

應用本體論於個人健康記錄系統之架構

The Architecture of Ontology-Based Personal Health Record System

王世羽
亞洲大學資訊科學
與應用學系
seeyeu@hotmail.com

林志豪
亞洲大學資訊科學
與應用學系
chlin@asia.edu.tw

摘要

病患參與決策越來越受到重視，促進醫病共同決策已成為 21 世紀醫療照護服務的目標之一。病患主動管理個人健康記錄有助於了解及溝通病情，但目前往往因為缺乏取的個人健康記錄的管道，和對專業名詞與醫療代碼不了解，降低了病患主動管理個人健康記錄的意願。本文目的在提供一個以病人觀點為主的應用本體論於個人健康紀錄系統架構。將各醫療單位 HIS 裡的病歷，透過 XML 格式的轉換，在資料庫中對照替換後，成為病患能了解的電子病歷紀錄，進而能更積極參與醫療計畫。
關鍵字：PHR、TMT、HL7、Ontology

Abstract

Patients participating in decision-making more and more pay attention. The promotion of medical decision-making has become one of the goals that 21st century medical care services. Patients with active management of personal health records contribute to the understanding disease and communication to doctor. Because of the Patients lack to access personal health records of the gateway, medical terminology and misunderstand the medical code. It is reducing the patient to actively manage their own health records will. This article aims to provide a patient-oriented view of the application of ontology in the personal health record system

architecture. The electronic medical record in HIS, through the translation of the XML format in the database were replaced and becoming patients easy to understand the electronic medical records .And Patients more actively involved in health care plan.

Keyword: PHR, TMT, HL7, Ontology

1. 前言

HL7、DICOM 及 TMT 等標準的製定，讓醫療資訊能在不同的醫療單位間互通有無，使醫療資源能夠更準確及有效的被利用。隨著國民知識水準的提升，社會大眾對醫療品質的要求日益增加，為了達到更完善的健康照護，需要醫病之間的良好互動；在西方國家，病患參與決策越來越受到重視，促進醫病共同決策已成為 21 世紀醫療照護服務的目標之一[11]。因此除了傳統被動接受治療行為，由病患擁有個人可攜式病歷以及病患主動管理個人健康記錄的觀念也漸漸盛行[26]。但目前往往因為缺乏取得個人健康記錄的管道，和對醫學名詞與代碼不了解，降低了病患主動管理個人健康記錄的意願。為解決此問題，提出應用本體論於個人健康醫療系統架構。本研究目的在於建置以病患觀點的個人健康資訊系統，應用本體論將病史作為一知識，讓病患能查詢自己的病歷狀況，並能了解其數值所代表的意義，進而對自身的健康管理有更深一層的認識及參與。

在醫病共同決策中，進行診療時醫生必須考慮病人的選擇與價值觀，以及醫生已不再是

病人唯一的資訊來源。隨著網際網路的快速發展，很多病人會從網路直接獲得資訊，並根據自己的需要和價值觀，進行分析及選擇[11]。但是從網路上的獲得的資訊，不一定符合自身病情的需要，故此本研究於系統架構內規劃建置自身醫療報告。系統導入 UMLS 正規化 HIS 所匯入醫療資訊，再將正規化後的資訊轉換成本體，透過代理人技術跟從醫療知識所匯入的相關的資訊與建立的病歷本體作資料對照；將比對結果代換原本文檔數值後，建立病歷報告輸出至使用者介面相對應功能區塊。

第 2 章探討本研究所應用的技術及標準，第 3 章為系統架構，第 4 章提出系統建置，第 5 章為結論，第 6 章為未來工作。

2. 文獻探討

2.1 Health Level 7

為了達到資訊共享的目的，將醫療資訊結構化、標準化，使其能透過網際網路，在不同醫事單位交換訊息。最多國家所使用的版本 HL7 2.3 版本，此版本被美國國家標準局(ANSI)通過認可為國家標準[7]。目前最新版本為 3.0，在架構上有相當大的變動，使用 Reference Information Model(RIM)，國際間仍處於學習和推廣的階段[8]。

2.2 Twain electronic Medical record Template(TMT)

台灣電子病歷模組(Taiwan Electronic Medical Record Template, TMT)，行政院衛生署有鑑於電子病歷將成為醫療機構未來病歷管理的發展趨勢，交由台灣醫學資訊學會協助制定及推廣電子病歷內容的基本格式，並應用國際醫療資訊標準，藉此統一全國電子病歷內容的基本格式，使醫療院所及產業有所遵循，有利於電子病歷分享與交換，提升醫療照護品質。

TMT 電子病歷格式為一個文件導向式的內容規範，與一般的傳統電子病歷資料庫不同，以 XML 技術為基礎的電子病歷文件規範，採用了 HL7 CDA V2.0 版本。TMT 格式與內容以醫療人員工作的紙本為基本單位，採用 XML 為架構，構成一個完整的樹狀之電子病歷架構。

衛生署委託資策會進行「國民健康資訊建設計畫(National Health Informatics Project, NHIP)」項目包括有將病歷保管的主控權，由醫院轉移給就診民眾本身，讓民眾取回自己健康資訊的所有權，並決定是否交由特定的健康資料銀行保管。因此 TMT 專案的目標即在於制定電子病歷內容基本格式(含確立及推動醫學詞彙標準)規範，供全國醫療院所參考，並訂定階段性執行目標，協助推動電子病歷[26]。

2.3 個人健康紀錄(Personal Health Record)

世界衛生組織(World Health Organization)[18]提倡醫療機構應該以主動預防的方式取代以往被動式的提供病患照顧，讓病患本身積極參與自身的健康照護。根據相關研究結果顯示，個人健康紀錄(Personal Health Record)的建立能讓病患對自己的病情更加了解，增加病患照護的滿意度，且讓病患更投入且配合治療計畫。PHR 包含各種不同的內容，但大致可分為以下項目[9]：

- (1) 基本資訊。
- (2) 過敏與藥物不良反應。
- (3) 用藥資訊，包括劑量及使用方式。
- (4) 疾病及住院資訊。
- (5) 手術及特殊治療。
- (6) 接種資訊。
- (7) 實驗數據及結果。
- (8) 家族病史。

美國健康資訊管理學會(AHIMA)將 PHR 定義為個人管理、隨處可得的終生健康資訊[1]。具備以下特性：

- (1) 個人可以掌管自己 PHR 的權限。
- (2) 包含個人終生健康資訊。
- (3) 包含來自醫療提供者的資訊。
- (4) 取得不受時間地點限制。
- (5) 資料隱密且安全。
- (6) 使用紀錄清楚明白，可以查詢何人何時何地使用過該 PHR 資料。

2.4 Extensible Markup Language

全球資訊網聯盟(W3C)在 1998 年所提出的一種置標語言，前身為 SGML(The Standard Generalized Markup Language)，由於 SGML 難以理解學習及 HTML 有擴充性和易讀性不佳等缺點，W3C 因而發展出 XML。[15]

2.4.1 XML

XML 有結構性、擴充性、驗證性與可讀性等特點存在，如表 1 所示。

表 1 XML 特點

特性	說明
結構性	巢狀式結構化語言
擴充性	元語言，沒有固定標籤，需要新標籤時，可以自行擴充定義。
驗證性	包含驗證內容的文件與文法；可以自行定義、下載文件型態定義(DTD)或 XML schema 驗證所編輯的 XML 文檔是否正確。
可讀性	文件結構是由許多元素(Element)所構成，而各個元素具有階層結構的特性，使用者可以自行選擇需要的部分判讀。

在 HL7 版本發展的過程之中，定義了資料內容和傳送順序，XML 是一套在網路上進行資料交換的語法標準；二者之間有著相當契合的互補性，如果將二者結合，對醫療資訊的交換有著顯著的幫助。所以在 HL7 2.5 版中加入了 XML 的編碼規則，並在 3.0 版中正式支援 XML[22]。

2.4.2 Extensible Stylesheet Language

Transformations

延伸樣式語言轉換(XSLT)，對 XML 文檔進行轉化的語言[21]，可將 XML 檔案轉換為另一種 XML 檔案，可分兩種轉換模式：

- Template-driven model：XSLT 樣板在 XML 文檔中尋找所需要的片段，並形成合乎 XSLT 要求的新文件。
- Data-driven model：在原始 XML 檔案中尋找匹配於 XSLT 的片段，形成類似原始文檔的新 XML 文件。

2.5 Ontology

本體論源於哲學，描述真實世界存在的實體並做有系統的說明。本體論廣泛應用於電腦科學的各個領域中，包括人工智慧、自然語言、資訊檢索和知識管理等等。不同領域對於本體論有不同的用途，學者對於本體論從不同角度提出定義。最常被引用的是 1993 年 Gruber 提出的定義：「An ontology is a formal, explicit specification of a shared conceptualization.」，本體論是對於群體共享的概念化之正式的、明確的表示形式。Conceptualization 指對領域確定現象的概念抽象模型；Shared 指本體論是共享的，屬於群體而非個人；Formal 指本體論是機器可讀的、可理解的；Explicit 指本體論的概念及限制用明確的方式表示出來[25]。

本體論用來描述特定領域下的知識時，可視為是概念、屬性、實例和關係的組合，以下分別說明這些元素：

- (1) 概念(Concept、Object、Class)：概念就是多個底層物件組合而成，也就是多個字彙(Vocabulary)的集合，該集合能用作概念的描述，描述主題的基本範圍。
- (2) 屬性(Attribute、Property、Slot、Role)：使用本體論表達知識領域時，概念是其子集合，將集合當作物件的話，物件間會有各種關係存在，且每一個物件擁有多種屬

性，屬性是用來描述物件的特徵。

- (3) 實例(Instance)：實例可用來表達上層概念，與上層概念存有某種關係，且繼承上層概念的屬性。常見一個建構本體論的問題，『某個項目要如何定義它是概念或是實例』，Kupiec 等指出，實例是用來更清楚的表達概念，因此在本體論的最底層定義出實例。
- (4) 關係(Relation)：本體論使用物件和屬性來清楚描述領域知識之外，仍不足以提供物件之間相關的細節訊息。因此，建構清楚的物件與屬性的本體論架構後，仍需定義物件間的關係。

2.5.1 TOVE 本體論工程

TOVE 本體論工程將本體論的建置分為六個階段[25]，如圖 1 所示

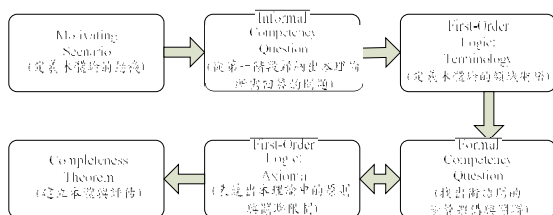


圖 1 TOVE 本體論工程

TOVE 本體論工程處理步驟說明：

1. 定義本體論的動機：此階段定義本體論的動機，同時描述所需面臨的問題，以及須解決的問題，和可能的解決方案。這個階段主要目的為了解本體的用途，幫助後續本體的維護動作。
2. 從第一階段中歸納出本體論所須回答的問題：從第一階段中歸納出本體論所須回答的問題從第一階段中歸納出本體論所須回答的問題，描述本體的需求，同時將藉此評估出是否符合使用者的要求。
3. 定義本體論的領域術語：此階段定義本體論會使用到的詞彙和詞彙之間的關聯性，也就是定義本體論的領域術語。TOVE 採用一階邏輯 (First-order logic) 來表達詞彙和關聯性。

4. 找出術語之間的階層架構與關聯：將第二階段的問題轉成標準的表示式，並找出術語之間的階層架構與關聯。在此階段所建立的標準表示式將成為評估本體的準則。
5. 表達出本體論中的原則與限制關係：使用一階邏輯表達出本體論中的原則、公理 (Axiom) 與限制、邏輯關係。
6. 建立本體與評估：建置本體，並評估所建立的本體是否符合第一階段的動機，以及能解決第二階段本理論所需面對的問題。

2.5.2 RDF

RDF 稱之為資源描述框架 (Resource description framework, RDF) [25]。人工智慧的術語裡，知識表達法(Knowledge representations) 可粗分為數理邏輯(Predicate calculus)、法則式專家系統(Rule-based expert system)、框架(Frames)及語意網(Semantic network)四大類[25]。而網際網路標準組織採用的資源描述框架就具有混合框架與語意網的特色，這裡知識是以物件 (Object)、屬性 (Attribute) 與值 (Value) 的三元架構來表達並可用有序對、有向圖及 XML 三種方式來表現。前兩者為闡述觀念，只有後者可以程式實作。

W3C 提出了 RDF 標準用來表達資源之間的關聯，希望能夠利用此一共通的資料通訊協定標準，讓網路環境中不同系統能夠互通資料。它是一個可以攜帶多種不同的元資料 (Metadata) 透過網路傳送與交換的工具。RDF 為資源這種資料模型提供了一種簡單的語義。資源描述框架基本以述語 (Predicate)、主詞 (Subject) 與受詞 (Object) 三個述句 (statement) 構成，其中述語代表屬性 (Property)，主詞代表資源 (Resource) 並以表達，受詞可為資源或文字 (Literal)。例如：現在有以籃球 (basketball) 作為受詞，其擁有者是 James，以有向圖來表現，其中橢圓代表主詞及受詞，箭號代表述語。如圖 2 所示。

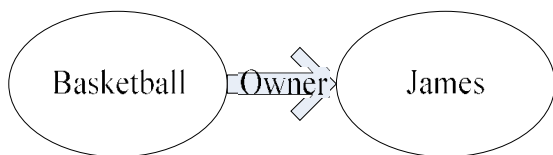


圖 2 RDF 表示法

RDF 的基本資料模式包含了三種物件型態：資源 (Resource)、特性 (Property) 與敘述式 (Statement)，並利用其三種物件型態來描述事物。RDF 沒有提供詳細的定義來描述一個資源所擁有的屬性，以及這些屬性與其他資源的關係。RDF Schema 則是用來表示資源的屬性定義與資源之間的定義，也可表示資源的類別、類別之間的定義與限制條件。

2.6 UMLS

美國國立醫學圖書館 [14] 研製了 UMLS(Unified Medical Language System)，運用類似自然語言處理的方法，針對生物醫學相關資訊中具有相同概念，但不同表達方式的項目，做關聯性的處理。UMLS 設立的目標，是電腦化且可持續發展的生物醫學檢索及情報資源指南系統。目的是能理解生物醫學的相關詞彙並幫助使用者檢索及獲取相關情報。除了基本的語言翻譯、自然語言處理、語言規範化的功能外，更重要的是能做跨資料庫檢索的詞彙轉換系統。連接包括電子病歷、臨床數據資料庫、專家系統...等資訊源，幫助使用者在不同資料庫間，做一致性的檢索。如此，可以克服不同系統語言差異性和分散資料庫的檢索問題 [3]。

UMLS 由四個部份所組成，

- (1) Metathesaurus
- (2) Semantic Network
- (3) Information sources Map
- (4) Specialist Lexicon

以下針對各部份做說明介紹：

2.6.1 Metathesaurus：

Metathesaurus 是一個關於生物醫學概念、

術語、詞彙等範疇的廣大集成索引 [13]。Metathesaurus 自發展開始即不斷擴充，2002 年第十三版中有約 210 萬個詞表達 776940 個概念，這些概念和詞來自 60 多個生物醫學詞典、術語典、分類表等引典 [13]。

Metathesaurus 是依據概念 (concept) 組織起來的，串接各個引典間概念與詞彙的關係，將相同概念的交替名稱和不同形式聯繫在一起，整合所有來源的索引典成為一個龐大的知識庫。

Metathesaurus 採用三級結構模式：Concept、Term 及 String，關係如圖 3 所示。單純形式相同的 String 被指向為同一 Term，而含義相同的 Term 又被 Concept 包含，不同的 Concept 間以 Relationship 相連。

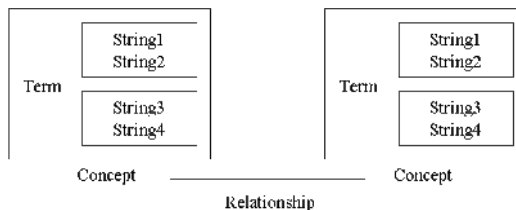


圖 3 三級結構模式

在 Metathesaurus 中，每一個概念被一個 CUI(unique concept identifier) 所定義；而個別 String，不管是大小寫還是語文上的拼字，只要有差異皆被指定一個 SUI(unique string identifier)；而連接概念與字串的 Term 就由 LUI(common term identifier) 所定義，也就是說 CUI 與 SUI 藉由 LUI 相連結 [13]。將 CUI、SUI 與 LUI 間的關係舉例說明，如表 2 所示。

字串 Atrial Fibrillation 與 Atrial Fibrillations 被認定為不同 String，給予不同的 SUI(S0016668、S0016669)，但這二字事實上只有單複數形式的不同，所以被指向同一個 LUI(L0004238)；同理 Auricular Fibrillation 與 Auricular Fibrillations 也連接至同一 LUI(L0004327)；但其實這四個字指的都是醫學上的「心房顫動」，所以二個 LUI 又被引領至同一 CUI(C0004238)。

表 2 CUI、SUI 與 LUI 間的關係

Concept (CUI)	Terms (LUI)	Strings (SUI)
C0004238	L0004238	S0016668
Atrial Fibrillation	Atrial Fibrillation	Atrial Fibrillation
Atrial Fibrillations	Atrial Fibrillations	S0016669
Auricular Fibrillation	Auricular Fibrillation	Atrial Fibrillations
Auricular Fibrillations	L0004327	S0016899
	Auricular Fibrillation	Auricular Fibrillation
	Auricular Fibrillations	S0016900
		Auricular Fibrillations (plural variant)

Metathesaurus 的代表性運用為 SNOMED (Systemized Nomenclature of Medicine), 其發展已超過 20 年, 內容包括: 解剖學、形態學、症狀及疾病體症、化學製品、藥品、有機體、物理因素、空間關係及職業等類別。目前諸多系統(如 Mesh)及臨床術語多選自本系統, 1999 年美國 30 多家單位開始測試使用本系統, 並持續提供意見以供系統改進。

2.6.2 Semantic Network

將各索引表間的概念以關係互相連接, 串接起來成為巨大的網狀架構, 即為語意網。而將不同索引間相同的概念串聯, 需要用到關係移轉的概念。所謂關係移轉, 意指若 A 與 B 同義, B 與 C 同義, 則可以說 A 與 C 同義, 如圖 4 所示。

將概念關係當成節點, 互相經由鏈(Link)連接起來即為語意網路。高層語義類型之間的關係, 可透過等級結構(鏈結)繼承給子類別, 但某些情況下直接繼承會令語義類型與關係的連接產生歧義, 即會阻斷[24]。

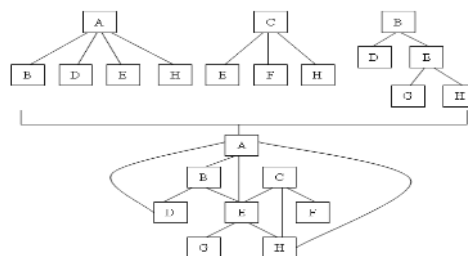


圖 4 關係轉移

2.6.3 Information sources Map

大量而快速成長的生物醫學資訊以及資料型態的多元化, 使得任何個人想要完整蒐集單一主題的生物醫學資訊, 十分困難。ISM 設計的目的即在幫助生物醫學領域的讀者, 經由系統功能之支援, 輔助其在資料內容上做選擇判斷, 並透過網路快速獲得相關資訊, 以解決其問題[24]。

2.6.4 Specialist Lexicon

專家詞語錄主要是在提供各項語詞資料, 以便系統可以據以處理自然語言所產生之語法上不確定的問題。我們可將 SPECIALIST Lexicon 視為是一套大部頭的電子辭典, 而其範圍則涵蓋一般常用英文單字及生物醫學辭彙[24]。也就是說, Specialist Lexicon 可以視為一套輔助詞典, 用來輔佐 Metathesaurus。

2.7 ICD-9

國際疾病分類(International Classification of Diseases, ICD)由世界衛生組織(World Health Organization)在歐洲早期對疾病編碼予以分類, 開始是用於疾病率和死亡率的統計, 之後逐漸變成針對疾病的國際標準的診斷分類及健康管理目標, 包含有人口健康狀況分析及疾病普及和影響的監控。台灣目前使用健保局所公佈之 ICD-9-CM 疾病碼一覽為基準[21]。

2.8 Agent

代理人(Agent)是指程式能依照事先使用者的指示, 偵測環境的變化並且作出相對應的行動。而代理人的兩個主要觀念為: 具有自主

性，並且為其目標而執行某些動作。代理人之間互相並存，並且共享資源，每個代理人有著其要扮演的角色，並且每個代理人只是代表一部份的功能[5]。設計用於代理使用者所明確指定的工作的一個實體代理人必須能獨立的工作，並能對事件有適當的反應及處理能力，且提供使用者溝通的介面，其特質[5]如下：

- (1) 自主性：能獨立的執行所指示任務，不須直接介入就能完成代理的工作。
- (2) 代理性：代理人能代表使用者完成其交付工作。
- (3) 溝通性：代理人與使用者透過特定介面或協定達到溝通的目的，接受使用者下達的指令或讓使用者知道某些資訊。
- (4) 監控性：代理人要能監控自己的環境，適時的調整本身的參數，以達成獨立作業。
- (5) 動作性：代理人要能自行感測環境，並對必須處理的事件作出反應。
- (6) 智慧性：代理人能對不同的事件作出適當的決策。

使用的目的不同，代理人可以分成使用者代理人(User Agents)、引導式代理人(Guide Agents)、自主式代理人(Autonomous Agents)、共存與合作式代理人(Symbiotic and Cooperative Agents)、擬人化代理人(Anthropomorphic Agents)等幾類[2]，將各分類整理說明如下：

- (1) 使用者代理人(User Agents)：代理人所執行的任務或工作都是使用者自己就可執行，只不過利用電腦可以提升效率及方便性。
- (2) 導引式代理人(Guide Agents)：像是引導者或老師，可以充實或補充使用者的知識，提供使用者他們需要的技術或專業知識。
- (3) 自主式代理人(Autonomous Agents)：行動不需要與使用者有互動的關係或是使用者輸入資料，可以自動且持續執行任務。
- (4) 共存與合作式代理人(Symbiotic and Cooperative Agents)：主要是提供使用者在

工作時一些額外的相關資訊以及從不同的角度來看這些工作。

- (5) 擬人化代理人(Anthropomorphic Agents)：模仿人類行為的代理人，如對話機器人。

2.9 RSA 演算法

RSA 演算法利用稱為公鑰(public key)及私鑰(private key)的二個質數來作加解密，公鑰用來對訊息加密，私鑰則用來解密，只要私鑰不公開，就算擁有公鑰也無法破解密文。RSA 演算法的可靠性基於分解極大的質數是相當困難的，只要鑰匙長度夠長，RSA 加密的訊息便不可能被暴力破解，如此可確保訊息安全[10]。

3. 系統架構

本研究建置應用本體論於個人醫療紀錄系統；將分散各家醫療單位的電子病歷整合，在系統內建置成個人病歷本體(Personal Health Ontology)整合病歷資訊，並透過 IT(information technology)技術擷取出內涵的資訊，且提供使用者查閱以了解自身的病史，架構如圖 5 所示。

以網路平台(Web Base)方式架構提供使用者使用上的便利性；各家醫療單位將使用者的 HIS 資料上傳至系統，系統可辨識 HL7 訊息格式及台灣電子病歷(TMT)訊息格式，進行訊息剖析，將資料轉換成 PHR 資料庫所需的 XML 檔案後，依其訊息格式輸入 XML 資料庫中對應欄位；XML 資料庫裡的可經由剖析轉換成 RDF(Resource description framework)，RDF 為建置本體的基本元素，以 RDF 導入預設的本體架構內建立個人病歷本體。PHR 資料庫導入 UMLS 技術，以泛索引典(Metathesaurus)以及語意網路(Semantic Network)和 ICD-9 的應用，正規化個人病歷中使用的名詞和病歷編碼；且匯入醫療知識(如藥典、基本數值參照表)協同建立診療報告。由於個人醫療紀錄是十分私密的個人隱私訊息，系統提供 RSA 加密機制以免個人資料遭竊取。

本系統運作流程可分為五部份，以下簡述：

1. HIS 資料轉換：系統將各家醫療單位的 HIS 資料轉換成 PHR 資料庫所需的 XML 檔案，存放至 XML Database 中。
2. 醫療知識轉換：透過代理人抓取醫療知識資料，依抓取資訊的格式不同，存取於資料庫、XML 資料庫或本體資料庫。
3. PHR 本體建立：依其轉入數據格式在 PHR 資料庫中轉換成 RDF；RDF 經由映對 (ontology mapping) 將 RDF 轉化成預設的本體架構中，以建立個人病歷本體；本系統設置本體伺服器查詢代理人 (Ontology Server Query Agent, OSQA) 以及本體流程管理代理人 (Ontology Process Manager Agent) 執行本體查詢與本體管理動作。本體存於本體資料庫 (Ontology Database)。
4. 診療報告建立：病歷本體與匯入資訊，將比對結果代換原本文檔數值後，建立單項病歷報告輸出至使用者介面相對應功能區塊。
5. 資訊安全；使用者登入 PHR 系統由 RSA 演算法加密確保資訊安全，登入後可獲取個人健康紀錄的相關資訊。

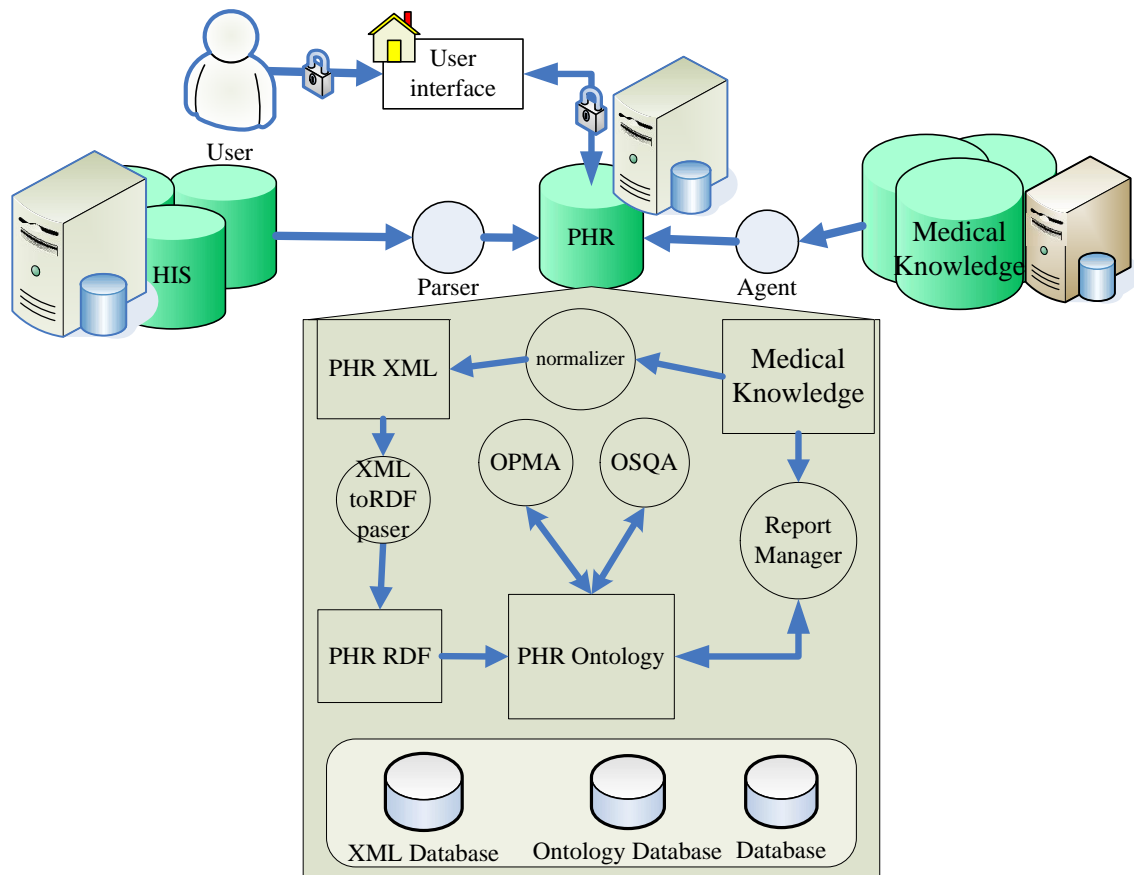


圖 5 系統架構圖

4. 系統建置

系統建置可分為這五部分：

1. HIS 資料轉換。
 2. 醫療知識轉換。
 3. PHR 本體建立。
 4. 診療報告建立。
 5. 資訊安全
- 分別說明如下。

4.1 HIS 資料轉換

國際間通用的標準格式健康資訊交換第七層協定 HL7，國內方面在衛生署健保局的指導下推廣已行之有年；近幾年來衛生署推動「網路健康服務推動計畫」[26]為了進行台灣電子病歷的交換與共享推行台灣電子病歷交換格式(TMT)。本研究採用 HL7 標準格式與台灣電子病歷交換格式為轉換格式。首先將各家醫療院所 HIS 中的 HL7 格式檔案與 TMT 格式檔案轉換為 PHR 資料庫所需的 XML 格式。HL7 格式版本採用 HL7 2.3 版本；資料轉換流程如圖 6 所示：

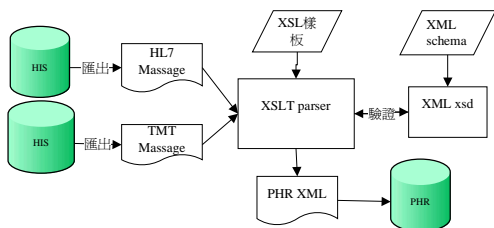


圖 6 HIS 資料轉換

由於各醫療院所 HIS 設計方式不同，但知道其各醫事單位 HIS 系統使用共同訊息格式，HL7 或 TMT，因而使用逆向工程[17]的方法；預先剖析 HL7 及 TMT 的訊息格式做出樣板及驗證碼。

第一步接收各醫療院所使用格式之 HL7 訊息及 TMT 訊息。第二步為 XML 剖析語法轉換，編輯吻合 PHR 資料庫格式的 XSL 樣板，在接收 HIS 系統中的 XML 格式後，按照樣板作 XSLT，產生 PHR 所需之 PHR XML 格式。

第三步為驗證 XML 檔案用，對 PHR 的 XML 格式編輯 XML schema 的 xsd 檔，利用該檔案來驗證所轉出的 PHR XML 檔案。本系統將各醫療院所 HIS 系統中不同格式轉換成 PHR 所需的 XML 格式

4.2 醫療知識轉換

由於建立醫療訊息運用代理人程式來抓取網際網路上的醫療知識，以建立 PHR 系統內個人醫療資訊的代碼、編碼及數值等的資訊來源，如圖 7 圖 7 所示。

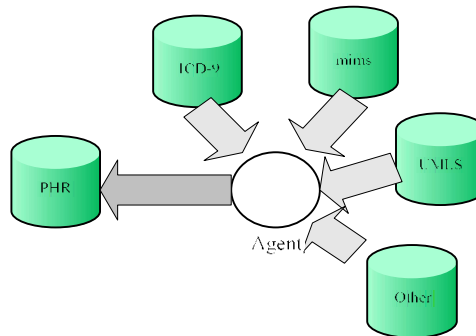


圖 7 醫療知識轉換

匯入醫療知識作為 PHR 病歷資訊的參照資訊，以下是匯入資訊的簡述：

- 基本數值參照表，可將 BMI 值、肝指數、血糖值、血壓等數值作對照比較，藉以讓使用者了解自己是否有過胖、過瘦、肝指數過高、高血壓等疾病，並可監視控制自身健康狀況。以 BMI 值為範例如表 3 所示；以血壓為例如表 4。

BMI	檢測結果	BMI	檢測結果
BMI<18.5	過輕	$27 \leq \text{BMI} < 30$	肥胖
$18.5 \leq \text{BMI} < 24$	正常	$30 \leq \text{BMI} < 35$	中度肥胖
$24 \leq \text{BMI} < 27$	過重	$\text{BMI} \geq 35$	重度肥胖

血壓分類	收縮壓 mmHg		舒張壓 mmHg
正常	<120	和	<80
高血壓前期	120-139	或	80-89
第一期高血壓 (輕度)	140-159	或	90-99
第二期高血壓 (中、重度)	≥160	或	≥100

- ICD-9 醫事編碼對照表[20]，將醫事編號代表何種病症表示出。使用健保局所公佈之 ICD-9-CM 疾病碼一覽為基準，建置 ICD-code 的英文、中文對照表，如表 5 所示。
- MIMS 藥典[10]，將藥物專有名詞作轉換，讓使用者了解自己所使用的藥物，提供用藥資訊，包括適應症狀、副作用及用藥需知。如圖 8 所示。
- UMLS 是電腦化且可持續發展的生物醫學檢索及情報資源指南系統。目的是能理解生物醫學的相關詞彙並幫助使用者檢索及獲取相關情報。

ICD code	英文	中文
001	Cholera	霍亂
002.0	Typhoid fever	傷寒
003.0	Salmonella gastroenteritis	沙門菌腸胃炎
:		
277.1	porphyria	紫質症
277.5	Mucopolysaccharidoses	黏多醣症

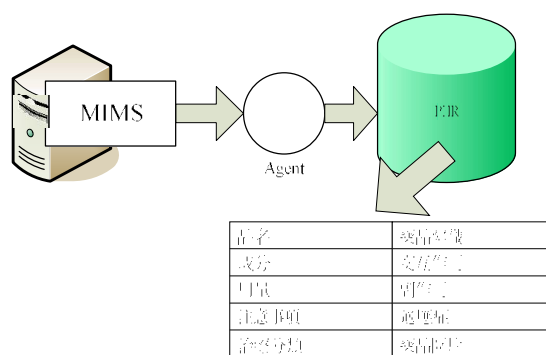


圖 8 藥典轉換

本系統使用代理人檢索及獲取相關情報；由於 UMLS 除了基本的語言翻譯、自然語言處理、語言正規化的功能外，還能做跨資料庫檢索的詞彙轉換系統。因此本系統設計代理人檢索 UMLS 時，由代理人匯入 UMLS 資料作為本系統的知識窗口，並透過知識專家來對所匯入的資料做驗證。如圖 9 所示。

本研究 PHR 資料庫導入 UMLS 技術，透過正規化代理人(Normalizer Agent)以泛索引典(Metathesaurus)以及語意網路(Semantic Network)和 ICD-9 的應用，正規化個人病歷中使用的名詞和病歷編碼。

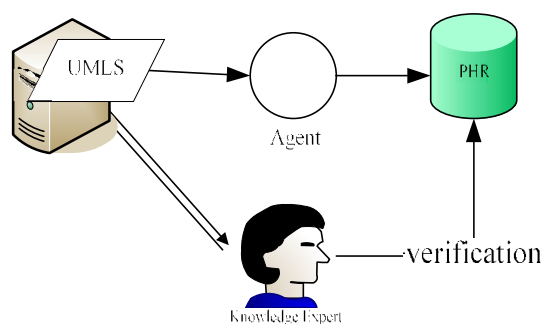


圖 9 UMLS 轉換

4.3 PHR 本體建立

本系統建立 PHR 本體以求將病史作為一知識，讓病患能查詢自己的病歷狀況。建立 PHR 本體第一步將 HIS 資料轉換進 PHR 的 PHR XML 檔案做轉換處理，此時 XML 檔案的輸入已經由正規化代理人(Normalizer Agent)正規化個人病歷中使用的名詞和病歷編碼，之後

再經由轉換代理人(XML to RDF parser)轉化成 RDF。RDF 為建置本體的基本元素，經由映對本體結構後，可建成 PHR 本體。第二步經由本體論映對(ontology mapping)將 RDF 轉化成預設的本體架構中；因各醫療院所 HIS 設計表單方式不同，但知道其各醫事單位 HIS 系統使用共同訊息格式，HL7 或 TMT，所以使用逆向工程[15]的方法，預先剖析 HL7 及 TMT 的訊息格式，依訊息模式製作出本體架構。但本體架構可能會有遺漏的部分需藉由校準本體流程回饋更新本體架構。建立使用使用者的病歷本體如圖 10 所示。

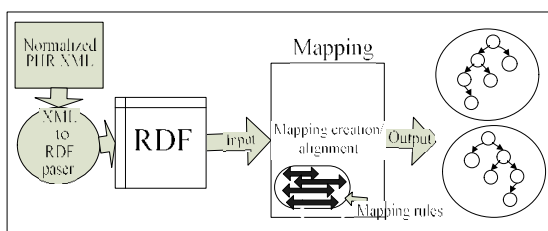


圖 10 RDF 轉換 ontology

HL7 ADT 轉換範例如圖 11 所示，已在剖析知道病歷的組成，ADT 由 MSH(訊息標頭)、EVN(事件類別)、PID(患者資料)、PV1(就醫資料)等區段所組成，MSH 訊息標頭內容包含訊息版本、訊息類型、發送機構、使用字元碼等；PID 患者資料內容包含性別、身分證、住址、電話等。以 MSH 與 PID 舉例映對規則(mapping rule)設計，像是 MSH、EVN、PID、PV1 等設計為概念(Concept)，MSH 所包含的訊息版本、訊息類型、發送機構、使用字元碼等就屬於實例(Instance)，MSH 與內含的訊息版本、訊息類型、發送機構、使用字元碼等的屬性關係(Attribute)即是包含(include)。

此外在本體控制與查詢部份，本系統設置本體伺服器查詢代理人(Ontology Server Query Agent, OSQA)以及本體流程管理代理人(Ontology Process Manager Agent)執行本體查詢與本體管理動作。

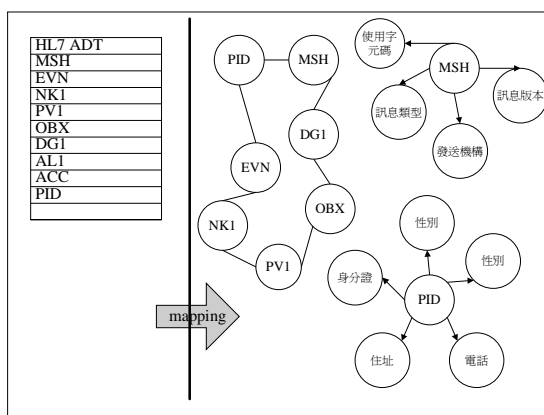


圖 11 HL7 ADT 轉換範例

建立單項本體並不是本系統所呈現 PHR 本體，需要透過校準本體流程(Ontology alignment process)將單項本體組合成病歷本體；流程如圖 12 所示。單項本體由於是從同一共同訊息格式 HL7 或是 TMT 編輯而成，因此其特徵相似度高，執行校準本體流程時特徵工程(Feature Engineering)特徵相同、相似點計算(Similarity Computation)相異度小與相似點聚集(Similarity Aggregation)，故容易呈現(Interpretation)且連結起網絡，但由於各醫事單位 HIS 表單設計的不同，組合病歷本體網絡時尚需要知識專家做本體結構及映對規則的調整。

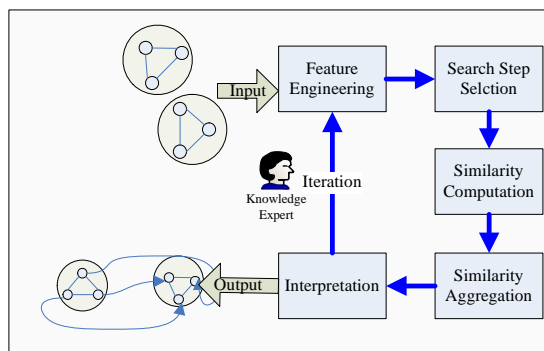


圖 12 Ontology alignment process

HL7 ADT 本體合併範例如圖 13 所示，ADT 本體由 MSH(訊息標頭)、EVN(事件類別)、PID(患者資料)、PV1(就醫資料)等區段所組成本體，MSH 訊息標頭內容包含訊息版本、訊息類型、發送機構、使用字元碼等組成本體；PID 患者資料內容包含性別、身分證、住

址、電話等組成本體。範例以 MSH 與 PID 透過校準本體流程加入至 ADT 本體內，MSH 本體、PID 本體與 ADT 本體結構上相異小，且有類似點可供聚集，顯示了使用同一共同訊息格式 HL7 或是 TMT 編輯而成的本體結構相容性高，進而代表本系統 PHR 本體與 HIS 原始病歷相異性也會相對的小。

藉由校準本體流程運行多次合併成病歷本體，建立個人病歷本體網絡，利用病歷本體網絡，在建置使用者泛索引典，將使用者的每項病歷當成是知識，使長期病歷本體網絡不受訊息版本變化(如 HL7 版本的改變 2X 到 3.0)的連續照護紀錄 (Continuity of Care Record, CCR) [5]。病歷本體網絡如圖 14 所示。

4.4 診療報告建立

隨著網際網路的快速發展，很多病人會從網路直接獲得資訊，並根據自己的需要和價值觀，進行分析及選擇。但是從網路上的獲得的資訊，不一定符合自身病情的需要，為解決此問題本研究於系統架構內規劃建置自身醫療報告。系統透過報告製作代理人 (Report Manager Agent) 跟從醫療知識所匯入的相關的資訊作資料對照，經過交叉查詢病歷本體，尋找本體領域 (Ontology Domain detection) 對應到關鍵特徵 (Keyword synonym matching) 及本體結構圖 (Ontology graph path)；病歷本體與匯入資訊，將比對結果代換原本文檔數值後，建立病歷報告輸出至使用者介面相對應功能區塊。如圖 15 所示。

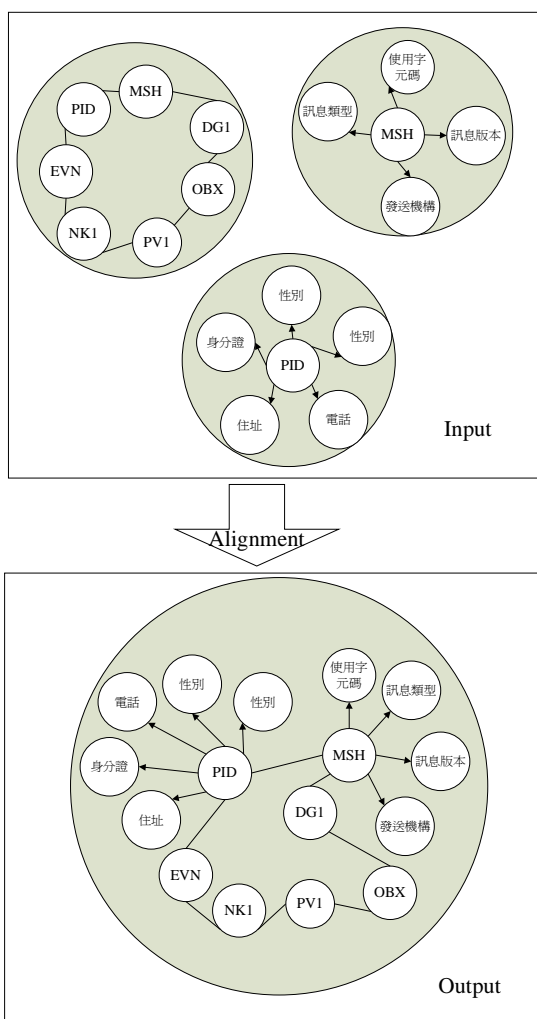


圖 13 HL7 ADT 本體合併範例

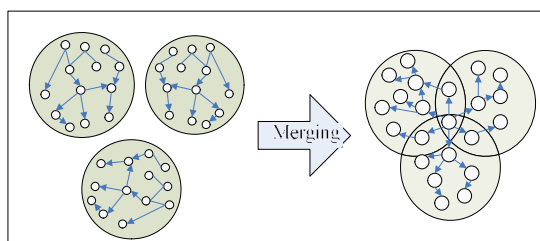


圖 14 病歷本體網絡

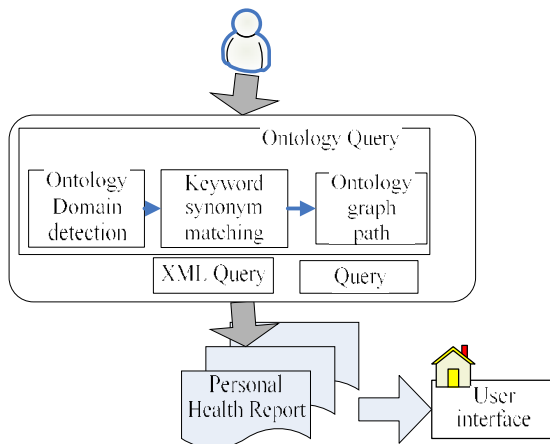


圖 15 Report process

4.5 資訊安全

使用者登入 PHR 系統使用者介面端(User interface)，確認使用者身份後，發送使用者欲查詢的要求給伺服器。PHR 伺服器到資料庫裡找查符合使用者需求的資料列表後進行加密動作，本研究使用 RSA 加密法將欲傳送資料加密成為密文，使用者從伺服器取得私密金鑰 (Private key)。PHR 伺服器接著將加密過後的文件送回介面端，由介面端回傳使用者，使用者再對密文進行解密，而獲得所查詢之資料。本安全架構以 RSA 的對稱式加密制度，確保使用者個人資料安全。如圖 16 所示。

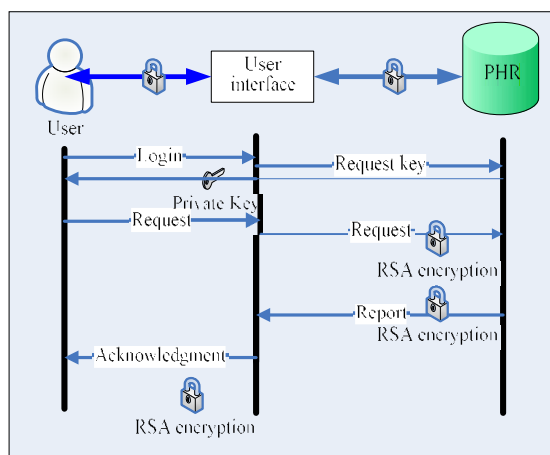


圖 16 資訊安全

5. 結論

HL7、DICOM 及 TMT 等標準的製定，讓醫療資訊能在不同的醫療單位間互通有無，使醫療資源能夠更準確及有效的被利用。除了醫療專業人員提供醫療行為，病患本身的也要積極參與才能達到完善的健康照護。以往缺乏取得個人健康記錄的管道，和對專業名詞與醫療代碼不了解，降低了病患主動管理個人健康記錄的意願。本研究提出應用本體論於個人健康醫療系統架構，系統導入 UMLS 正規化 HIS 所匯入醫療資訊，再將正規化後的資訊轉換成本體，透過代理人技術跟從醫療知識所匯入的相關的資訊與建立的病歷本體作資料對照，建立醫療報告使病患獲得符合自己病情的知識；建立病歷本體網絡是連續照護紀錄的概念，對一生的病歷做系統化的整理，不受 HL7、DICOM 與 TMT 等的醫療標準版本變化而帶來的衝擊，建立此系統建立病患能了解的電子病歷紀錄，進而能更積極參與醫療計畫。

6. 未來工作

本系統使用病歷為導向，病歷是經由醫事人員處理後的訊息，因此無法提供病人即時的身體訊息，僅能被動提供以往病歷史相關資訊提醒注意事項，未來將擴充系統功能，增加自然語言處理 Q&A 系統，提供病患如有相關病情上的問題能夠獲得最即時正確的答案或建議事項。

參考文獻

- [1] AHIMA e-HIM PHR Work Group, "Defining the Personal Health Record", *Journal of AHIMA/American Health Information Management Association*, 76, 24-25, 2005.
- [2] Beale, R. and Wood, A., "Agent-Based Interaction," *People and Computers IX: Proceedings of HCI'94, UK*, 1994.

- [3] Bodenreider, Olivier, "The Unified Medical Language System (UMLS): integrating biomedical terminology". *Nucleic Acids Research*, 32, D267-D270, 2004.
- [4] Browne, McCray and Srinivasan, "The Specialist Lexicon. Lister Hill National Center for Biomedical Communications", National Library of Medicine, Bethesda, MD, p. 1, 2000.
- [5] Eric A. Coleman, MD, MPH, and Robert A. Berenson, MD, "Lost in Transition: Challenges and Opportunities for Improving the Quality of Transitional Care", 2004[CCR]
- [6] Hayes, C., "Agents in a Nutshell - A Very Brief Introduction," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 11, No. 1, January/February 1999.
- [7] HL7, "Health Level 7 Version 2.3.1 Final Standard", 1999.
- [8] HL7, "HL7 Clinical Document Architecture, Release 2", 2006.
- [9] M. I. Kim and K. B. Johnson, "Personal Health Records: Evaluation of Functionality and Utility", *Journal of American Medical Informatics Association*, 9, 171-180, 2002.
- [10] MIMS, <http://www.mims.com/>
- [11] Muir Gray, 唐金陵, 實證醫學醫療照護決策, 2007
- [12] R. Rivest, A. Shamir and L. Adleman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems", *Communications of the ACM*, Vol.21, No.2, pp. 120-126, Feb 1978.
- [13] Smith, Barry Kumar, Anand and Schulze-Kremer, Steffen, "Revising the UMLS Semantic Network", in M. Fieschi, et al. (eds.), *Medinfo 2004*, Amsterdam: IOS Press, 1700, 2004
- [14] UMLS, <http://www.nlm.nih.gov/research/umls/>
- [15] W3C, "Extensible Markup Language", <http://www.w3.org/XML>.
- [16] W3C, RDF Vocabulary Description Language 1.0: RDF Schema, 10 February 2004, <http://www.w3.org/TR/rdf-schema/>
- [17] Warden, R. (1992) · *Software Reuse and Reverse Engineering in Practice · Software Reuse and Reverse Engineering in Practice* · London, England : Chapman & Hall , 283-305 ·
- [18] WHO, <http://www.who.int/en/>
- [19] WHO, International Classification of Diseases (ICD), <http://www.who.int/classifications/icd/en/>
- [20] 中央健康保險局, 疾病分類代碼及範圍, <http://www.nhi.gov.tw/>
- [21] 中央健康保險局, 2001 年 ICD-9-CM 疾病碼一覽表, <http://www.nhi.gov.tw>。
- [22] 林育民, 設計以 HL7/XML 為基礎的巡迴醫療資訊系統, 私立中原大學工業工程學系碩士論文。
- [23] 張智勛、黃衍文、吳季勳、李麗惠、劉德明, 產生 HL7 V3 訊息之研究: 以病患資料查詢為例, 第四屆亞太 HL7 健康資訊交換標準研討會。
- [24] 張慧銖, 統一醫學語言系統簡介, 台大醫學院圖書分館館訊第廿七期第 1-4 頁, 2007。
- [25] 郭子鼎, 應用本體論於國中社會科線上問答系統, 亞洲大學資訊科學與應用學系碩士論文。
- [26] 臺灣電子病歷交換基本格式, <http://emr.doh.gov.tw/old/index.html>