

The Construction of a travel information retrieval system on Semantic Web

葉木水、陳鴻文

E-mail: 9314357@mail.dyu.edu.tw

ABSTRACT

The issue of effectively using information and services become more important among Internet technologies. The concept of "Semantic Web" was thus proposed to make Internet documents re-usable and explicit based on "Ontology". This research focuses on two folds: the ways of defining and constructing a Taiwan travel ontology, and then an efficient mechanism of information retrieval on Semantic Web. Since there is no tourist ontology available on Internet now, relevant information for trip were collected to define and build domain and information ontologies. In order of an efficient indexing the constructed information ontologies were further translated into relational databases from original DAML format. Besides, "top-down" management of Internet services was adopted; In other words, each ontology and related services need to be recorded in the registration center in advance. Under the environment, a novel information retrieval system using tag hierarchy within ontology was proposed and compared with traditional search mechanism. Four experiments illustrated that the results of proposed system was more accurate and less redundant. Therefore, the now services based on Semantic Web could be expected to match users' demands more efficiently by offering the designed ontologies and proposed information retrieval mechanism.

Keywords : Semantic Web, Web Service, Information Retrieval, Ontology, DAML, JENA

Table of Contents

封面內頁 簽名頁 授權書 iii 中文摘要 v 英文摘要 vi 誌謝 vii 目錄 viii 圖目錄 x 表目錄 xii 第一章 緒論 1 第一節 研究背景與動機 1 第二節 研究目的 3 第三節 研究範圍與限制 4 第四節 研究內容與研究流程 5 第二章 文獻探討 7 第一節 語意網 7 第二節 語意網語言 9 第三節 知識本體 14 第四節 知識本體的開發工具 18 第五節 資訊檢索的探討 21 第六節 WSDL簡介 23 第三章 系統設計與研究方法 26 第一節 系統假設 26 第二節 都會旅遊知識本體 27 第三節 系統雛型架構 32 第四章 系統實作與效能評估 44 第一節 開發工具與環境 44 第二節 系統實作 44 第三節 檢索實驗與方法 53 第五章 結論及後續研究建議 61 第一節 研究結論 61 第二節 未來研究方向 63 參考文獻 64 附錄A：旅遊詞彙庫 69 附錄B：階層編號表 70 附錄C：交通工具資訊(1100)部份資料內容 71 附錄D：景點資訊(1200)部份資料內容 72 附錄E：住宿資訊(1300)部份資料內容 73 圖目錄 圖1.1 研究流程.....6 圖2.1 語意網上網頁設計語言之層級模型.....8 圖2.2 XML文件.....10 圖2.3 RDF範例.....11 圖2.4 Prot?g?-2000操作介面範例.....20 圖2.5 WSDL架構.....23 圖2.6 WSDL文件內容.....24 圖3.1 本論文所採行之知識本體涵蓋內容.....27 圖3.2 本論文採行之旅遊知識本體.....28 圖3.3 本論文所採行之交通領域知識本體.....29 圖3.4 本論文所採行之景點領域知識本體.....30 圖3.5 本論文所採行之住宿領域知識本體.....30 圖3.6 以DAML語法描述景點部份領域知識本體.....31 圖3.7 以DAML語法描述部份之景點資訊實體.....32 圖3.8 具有資訊檢索功能之旅遊語意網整體服務系統架構.....33 圖3.9 標準SOAP訊息的內容.....34 圖3.10 資訊檢索模組之運作流程範例.....35 圖3.11 都會旅遊資訊實體樹狀架構之階層編號.....37 圖3.12 階層編號表資料結構.....38 圖3.13 旅遊詞彙庫資料表內容.....39 圖3.14 資訊檢索系統之標籤檢索部份程式碼.....41 圖3.15 判定關鍵字詞階層關係的程式碼.....42 圖4.1 景點資訊類(木柵動物園)之部份DAML檔案內容.....45 圖4.2 「交通工具資訊」之資訊實體資料(DAML格式).....46 圖4.3 DAML檔與資料庫格式轉換的部份程式碼.....48 圖4.4 轉換程式執行結果.....49 圖4.5 資料表格實際轉換的內容.....49 圖4.6 資訊實體檢索系統之「樹狀結構查詢」的介面.....50 圖4.7 資訊實體檢索系統之「樹狀結構查詢」功能展開的介面.....51 圖4.8 資訊實體檢索系統之「一般查詢」介面.....53 圖4.9 樹狀結構查詢之「查詢一」檢索結果.....55 圖4.10 「樹狀結構查詢」之「查詢二」檢索結果.....56 圖4.11 一般查詢之查詢一檢索結果畫面.....58 圖4.12 一般查詢之查詢二檢索結果畫面.....60 表目錄 表2.1 PROT?G? 2000與KAON的比較.....19 表4.1 樹狀結構查詢之查詢一條屬性表(至第二層).....54 表4.2 樹狀結構查詢之查詢二條件屬性表(至第三層).....56 表4.3 一般查詢之查詢一檢索結果的比較.....58 表4.4 一般查詢之查詢二檢索結果的比較.....60
--

REFERENCES

1.顧值豪, 跨網站式XML購物引擎之建構 - 以手機選購為例, 大葉大學資訊管理所碩士論文, (2001)。 2.魏召欣, 智慧型企業文件檢索系統之建置 - 以大葉大學公文檢 索為例, 大葉大學資訊管理所碩士論文, (2003)。 3.高嘉祺, 線上圖文購物引擎-以手機應用為例, 大葉大學資訊管 理研究所碩士論文, (2000)。 4.朱毓君, 以本體論強化網路FAQ系統之解答整合能力, 國立台灣 科技大學電子工程系碩士論 文, (2001)。 5.陳詩沛, 全球資訊網上的語意搜尋, 國立台灣大學資訊工程學研 究所碩士論文, (2002) 6.Tim Berners-Lee, James handler, Ora Lassila撰文, 高虹 譯, "【資訊科技】電腦也能看懂", 科學人, p47-p56, (2002)。 7.郭尚君,郭慧琦編著, "XML技術實務", 文魁資 訊股份有限公司, 初版1刷, (2002)。 8.Jerome Euzenat, INRIA Rhone-Alpes, James Hendler 編輯, -"Research Challenges and Perspectives of the Semantic W -eb", IEEE Intelligent systems, p86-p88, (2002)。 9.Dieter Fensel and Frank van Harmelen, Ian Horrocks, Debo -rah L. McGuinness, Peter F. Patel-Schneider, "OIL:An O -ntology Infrastructure for the Semantic Web", IEEE Inte -lligent systems, p38-p45, (2001)。 10.Natalya F. Noy and Deborah L. McGuinness, "Ontology Dev -elopment 101: A Guide to Creating Your First Ontology" -, <http://protege.stanford.edu/useit.html>, p1-p25。 11.S. Staab and A. Maedche, "Knowledge Portals Ontologies -at Work", AI Magazine, Vol. 22, No. 2, p63-p75, (2001)。 12.B. Chandrasekaran, J. R. Josephson, and V. R. Benjamin -s -, "What Are Ontologies, and Why Do We Need Them?", IEE -E Intelligent Systems, p20-p26, (1999)。 13.C. Y. I. Lin and C. S. Ho, "A Generic-Ontology-Based Appr -oach for Requirement Analysis and its Application in Net -work Management Software", Artificial Intelligence for E -ngineering Design, Analysis and Manufacturing, Vol. 13, N -o. 1, p37-p61, (1999)。 14.N. Guarino, "Formal Ontology and Information Systems," Pro -c. Of the 1st International Conference on Formal Ontologi -es in Information Systems, FOIS'98, pp.3-15. Trento, Ital -y, Amsterdam, ISO Press, (1998)。 15.Asunci?n G?mez-P?rez and Oscar Corcho, "Ontology Languages - for the Semantic Web," IEEE Intelligent Systems, pp.54-60, (2002)。 16.R. Scott Cost, T. Finin and A. Joshi, "ITtalks: A Case Stud -y in the Semantic Web and DAML+OIL," IEEE Intelligent Systems, pp.40-47, (2002)。 17.Sheila A. McIlraith, Tran Cao Son, and Honglei Zeng, "Seman -tic Web Services," IEEE Intelligent Systems, pp.46-53, (2001)。 18.Ian Horrocks, "DAML+OIL: a Description Logic for the Semant -ic Web," IEEE Computer Society Technical Committee on Data Engineering, (2001)。 19.Deborah L.McGuinness, Richard Fikes, James Hendler, and Lynn - Andrea Stein, "DAML+OIL: An Ontology Language for the Seman -tic Web,"IEEE Intelligent Systems, p72-p80, (2002)。 20.Stefan Decker, Sergey Melnik, Frank Van Harmelen, Dieter Fens -el, Michel Klein, Jeen Broekstra, Michael Erdmann, and Ian Ho -rrocks, "The Semantic Web: The Roles of XML and RDF," IEEE In -ternet Computing, p63-p74, (2000)。 21.James Hendler, "Agents and the Semantic Web," IEEE Intelligent -Systems, p30-p37, (2001)。 22.Michael N. Huhns, and Munindar P. Singh, "Ontologies for Agents -," IEEE Internet Computing, p81-p83, (1997)。 23.Venu Vasudevan, "A Web Services Primer," <http://www.xml.com/pub> -/a/ws/2001/04/04/webservices/index.html, (2003)。 24.Michael Uschold, "Where are the semantics in the semantic web?" -, AI Magazine, p25-p35, (2003)。 25.Brian McBride,"Jena: A Semantic Web Toolkit", IEEE Internet Com -puting, p55-p59, (2002)。 26.Wolfgang May,"Linking the Semantic Web with Existing Sources", I -EEE Intelligent Systems, p1-p5, (2002)。 27.Isabel F. Cruz and Afsheen Rajendran, "Semantic Data Integratio -n in Hierarchical Domains", IEEE Intelligent Systems, p66-p73, (2003)。 28.W3C, <http://www.w3c.org>。 29.O'REILLY XML.com, <http://www.xml.com>。 30.Jena Semantic Web Framework. <http://jena.sourceforge.net/documentation.html> 31.Jena Tutorial for Release 1.4.0. <http://www.hpl.hp.com/semweb/doc/tutorial/> 32.Annotated DAML+OIL Ontology Markup. <http://www.w3.org/TR/daml+oil-walkthru/> 33.DAML.Org. <http://www.daml.org/> 34.台北市政府。 http://www.taipei.gov.tw/cgi-bin/classify/index.cgi?class_id=%41%30%34%2C%42%30%33 35.台中市政府。 <http://www.tccg.gov.tw> 36.台南市政府。 <http://www.tncg.gov.tw/01.asp?sub1=市府組織&sub2=交通局&sub3=交通管理課&page=1> 37.高雄市政府。 <http://www.kcg.gov.tw/> 38.中華民國交通部觀光局。 <http://taiwan.net.tw/lan/Cht/search/index.asp> 39.大台中觀光旅遊網。 <http://travel.tccg.gov.tw/index1.asp?Litem=C1> 40.高雄網。 <http://kaohsiungwalking.kcg.gov.tw/> 41.台灣電子地圖服務網。 <http://www.map.com.tw> 42.Google。 <http://www.google.com.tw> 43.昇陽(Sun)電腦公司。 <http://www.sun.com.tw>