

Career Guidance: A Semantic Web Approach

Dhanushka Anuradha Kodituwakku
08/10034



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfillment of the requirements of the Degree of Master of Science in Information Technology

University of Moratuwa



102488

February 2011

004 "11"
004 (043)

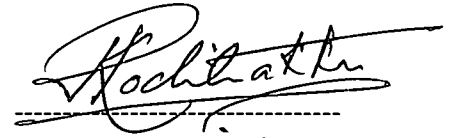
74

102488

Declaration

We declare that this thesis is our own work and has not been submitted in any form for another degree or diploma at any university or other institution of tertiary education. Information derived from the published or unpublished work of others has been acknowledged in the text and a list of references is given.

Dhanushka Anuradha Kodituwakku



Date: 09/12/2011



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Supervised by

Dr. G. D. S. P. Wimalaratne



Date: 9/12/2011

Dedication

I would like to dedicate this dissertation to my loving parents, Mr. Jayasena Kodituwakku and Mrs. Gladis Kodituwakku, and my ever supportive wife Nadishani Kodituwakku.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Acknowledgment

First and foremost, I would like to thank Dr. G. D. S. P. Wimalarathne, my supervisor, for the immense guidance and constructive suggestions given with ceaseless patience to complete this research successfully.

Furthermore, I would like to express my unfathomable sense of gratitude toward all who helped me in any way during the course of the research.

This work was conducted using the Protégé resource, which is supported by grant LM007885 from the United States National Library of Medicine.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk


Abstract


Choosing a career is one of the major decisions that are taken during the life span of a human. Its impact on the person and on the society is definite. Career information is the most essential input to the career decision making process. This information should be complete and well structured for successful decision making. The World Wide Web is the single most large information source which contains career information. However, this information is scattered through out it with no real structure. Semantic Web has emerged as a feasible solution for centralizing and structuring data. Therefore, this research aims at designing and creating a Semantic Web Application that can collect and organize career information effectively and efficiently. The application simply collects data, published in RDF/XML format, shared by various entities. The resulting RDF store is then queried using the SPARQL query language. The career ontology presented in the research paper is used to model the information portal and is a vital part of it. Researcher also stresses on how important the inference mechanism is for the Semantic Web application in organizing the content logically.



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Contents

	Page
<i>Declaration</i>	<i>i</i>
<i>Dedication</i>	<i>ii</i>
<i>Acknowledgement</i>	<i>iii</i>
<i>Abstract</i>	<i>iv</i>
<i>List of Figures</i>	<i>viii</i>
<i>List of Tables</i>	<i>x</i>
Chapter 1 – Introduction	1
1.1 Background and Motivation	1
1.2 Aim and Objectives	5
1.3 Proposed solution in brief	5
1.4 Structure of the dissertation	7
1.5 Research chapters	7
 University of Moratuwa, Sri Lanka. Electronic Theses & Dissertations www.lib.mrt.ac.lk	
Chapter 2 – Literature Review on Career Guidance Information & Semantic Web	8
2.1 Introduction	8
2.2 Importance of career guidance and career information	8
2.3 Economic Implications	12
2.4 Importance of career guidance in Sri Lanka	13
2.5 Linked data to centralize information	13
2.6 Summary	14
Chapter 3 – Centralize and Organize Career Information Using Semantic Web	15
3.1 Introduction	15
3.2 Congregate and organize career information	15
3.3 Semantic Web to the rescue	19

3.4 Summary	24
Chapter 4 – Technology Adapted - Semantic Web Approach	25
4.1 Introduction	25
4.2 Overview	25
4.3 Solution	26
4.3.1 Users and Inputs	26
4.3.2 Technology	29
4.3.3 Process	30
4.3.4 Output	31
4.4 Summary	31
Chapter 5 – Analysis and Design	32
5.1 Introduction	32
5.2 Methodology	32
5.3 Prototype system's modules	33
5.4 Summary	35
 University of Moratuwa, Sri Lanka. Electronic Theses & Dissertations www.lib.mrt.ac.lk	
Chapter 6 – Designing Web Ontology for Career Information Portal	
6.1 Introduction	36
6.2 Role of career ontology	36
6.3 Career ontology	37
6.4 Summary	43
Chapter 7 – Implementation	44
7.1 Introduction	44
7.2 Development Environment	44
7.3 Component Implementation	47
7.3.1 Login	47
7.3.2 User Manager	48
7.3.3 Site Manager	49
7.3.4 RDF Collector	51

7.3.5 Ontology Browser	52
7.3.6 RDF Browser	52
7.3.7 Search	53
7.3.8 Database Access	54
7.4 Summary	54
Chapter 8 – Testing and Evaluation	55
8.1 Introduction	55
8.2 Testing	55
8.2.1 Testing Plan	55
8.2.2 Test results evaluation	56
8.2.3 Testing conclusion	56
8.3 Evaluation	57
8.3.1 Evaluation Methodology	57
8.3.2 Heuristic Evaluation Results	58
8.3.3 User Evaluation Survey	59
8.4 Summary	60
Chapter 9 – Conclusion and Further work	61
9.1 Introduction	61
9.2 Accomplished Aims and Objectives	61
9.3 Problems encountered	62
9.4 Personal achievements	62
9.5 Research limitations	63
9.6 further work	63
9.7 Conclusion	63



List of Figures

	Page
Figure 1.1: Information volume and labor supply	3
Figure 1.2: Performance of Candidates	4
Figure 1.3: User using portal interface for information	6
Figure 2.1: A Career Decision Making Model	10
Figure 2.2: World Internet users and population statistics by sector and province – 2009	11
Figure 3.1: Career information input, processing and output	15
Figure 3.2: Making information complete	18
Figure 3.3: TIOBE programming community index for 2011	17
Figure 3.4: Combined data forming useful information	20
Figure 3.5: Data modeling methods	22
Figure 3.6: Real world example for Data Graph	22
Figure 3.7: Data graph to RDF/XML	23
Figure 4.1: User interface of proposed web portal – front page	27
Figure 4.2: User interface – information drill down	27
Figure 4.3: Data publishing via RDF	28
Figure 4.4: Information of the song Elevenpath by Nightwish available in RDF	29
Figure 5.1: Component diagram	33
Figure 5.2: URL Submit Form	34
Figure 6.1: Ontology Extends vocabulary	37
Figure 6.2: CareerPath class hierarchy	38
Figure 6.3: Sample instances of the Software Engineer career	39
Figure 6.4: RDF file with job vacancy data	40
Figure 6.5: Data graph for RDF data	40
Figure 6.6: Graph pattern for querying Software Engineer vacancies	41
Figure 6.7: Screen shot from CRP (Job vacancy listings)	42

Figure 7.1: Protégé – OWL Editor	46
Figure 7.2: Login Component	47
Figure 7.3: Sequence Diagram for Login	48
Figure 7.4: Site Manager Component	49
Figure 7.5: Sequence Diagram – URL Submission	50
Figure 7.6: URL approval interface	50
Figure 7.7: RDF Collector Component	51
Figure 7.8: RDF File collection process	51
Figure 7.9: Job vacancy retrieval from InfModel	53
Figure 7.10: Sample SPARQL query execution	53
Figure 8.1: Test Plan	55
Figure 8.2: Testing evaluation graph	56



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

List of Tables

	Page
Table 2.1: Internet Email using population (Percentage) 5 – 69 years aged	11
Table 3.1: Short comings of primary data sharing approaches	21
Table 6.1: Results for 6.5 graph pattern	45
Table 7.1: ASP.NET C# Semantic Web Tools	37
Table 8.1: Heuristics used in Heuristic evaluation	57
Table 8.2: Questionnaire result summary	58
Table 9.1: Accomplished research objectives and aim	61



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk