A Job Recruitment System Using Semantic Web Technology

P. Niaphruek

Department of Computer Science, Faculty of Science,
Rajamangala University of Technology Thanyaburi, Klong 6, Thanyaburi, Pathumthani 12110

E-mail: pongpon@gmail.com

Abstract

This research presented a new job recruitment system using Semantic Web technology. First, it introduced an information model to represent a CV (curriculum vitae) by creating the CV Ontology based on exiting well-known standards such as Europortfolio Europass CV, Description of a Career, FOAF and VIVO. Second, the research designed a new business model for a job recruitment system using benefits of the proposed CV Ontology. Finally, the research implemented a prototype system based on the proposed business model and evaluated it against other popular online recruitment systems in order to prove the feasibility of the approach. In sum, the result expresses that this proposed system is quite better than others, especially in supporting RDF, supporting ontologies and vocabularies, supporting SPARQL, and enabling linked data by SPARQL endpoint. These can solve the lack of a well-defined meaning of information and enable information-sharing across job portals.

Key Words: Job recruitment system, Implementation, Information Model, Linked Open Data, Ontology, Semantic Web technology, RDF, OWL.

1. INTRODUCTION

During the past decade, Internet access has interestingly risen from zero to seventy percent of the population and this rapid increase has the potential to vividly alter labor market outcomes (Stevenson, 2008). Furthermore, the increase of a global Internet culture during this time has influenced the way that most recruiters have a tendency to collect information of applicants from the Internet. An online recruitment system presently has been used as a tool to recruit appropriate employer since it is the process of personnel recruitment which uses electronic resources particularly the Internet (Stone, 2006). Apart from that there are many people who can access the Internet everywhere by using various devices such as a laptop, a smart phone, and a tablet; they normally use Internet to post their personal information. According to that, companies and recruitment agencies have quickly changed their common recruitment process into online method with the aims at improving the speed of matching the applicants and job requirement.

In the area of job recruitment, the trend is in the direction of online exchange of information about human resources. Nowadays, there are many people widely using a job portal to recruit a set of desired candidates or jobs. However, many unqualified applicants may be generated by online recruitment system as well as it is still difficult to find the best personnel (Robbins & Coulter, 2009). Furthermore, to share the information across job portals is quite problematic because it has various formats. Likewise, applicants' information still lacks the well-defined meaning.

In order to make a well-defined meaning of information, Semantic Web technology is strongly recommended to increase meanings of information, which can enhance the matching between requirements and information. Besides, it can also enable information-sharing among web portals and reduce the duplicated information.

In this paper, it aimed to design an information model for representing the applicant's information which is commonly named CV (curriculum vitae or resume). Then, to design a new business model for job application and personnel recruitment in which each applicant manages his/her own personal information and in which a recruiter can directly access an applicant's personal information. After that, to develop a prototype of a recruitment system which is based on the proposed business model and information model. The next is to evaluate it against other popular online recruitment systems to prove the feasibility of the approach. Finally, the conclusion and future work are described in the final section.

2. RELATED WORKS

Nowadays, the instances of the best job search engines include Craigslist[1], CareerBuilder[2], Indeed[3], and Monster[4] (List of Search Engines, 2012) which encounter with various problems such as finding a suitable person, sharing the information among web portals, and lack of a standard for keeping the information. There are various researches that attempted to solve the mentioned problems by using several techniques, especially in the Semantic Web Technologies. Bizer, Heese, Mochol, Oldakowski, and Tolksdorf described how online recruitment processes can be efficient using Semantic Web technologies. They analyzed the understanding chances of such a situation from the technological and business viewpoint. Also, they described a typical implementation of the essential technological foundation and analyze the possible economic impacts, the interests of the complicated parties, and potential alteration to their business models (Bizer, Heese, Mochol, Oldakowski, & Tolksdorf, 2005). Dorn, Naz and Pichlmair tried to create an ontology which is deliberated by its incorporation of job descriptions, ideas for evaluating competences on different levels and confirmations for capabilities (Dorn, Naz, & Pichlmair, 2007). They suggested measuring the level of experience and knowledge by using the proficiency of expert such as a head of a department who is an expert in business; he/she can measure the level of knowledge up to the highest level. In contrast, professor at a university can measure the level of knowledge up to the maximum level. A prototype for integrating information of different job portals into one meta-search engine was proposed by Dorn and Naz also (Dorn & Naz, 2007). They used the HR-XML-schema to build a form for searching jobs. Moreover, the standard of DOAC[1] and Europass (European Commission Group, 2010) have been used to make the recommendation tagging for associated information about the computer skills in (Mirizzi, Noia, & Sciascio, 2009) so that it can semantically recommend related information.

Designing Information Model

In creating information models, this study only aimed to design ontology for storing information of an applicant, or it is called "CV ontology". According to the standards of CV, it was found that the existing CV ontologies such as ResumeRDF and DOAC are a proper CV ontology for use, but both are out-of-date and they have no continual improvement so far. Another interesting standard of CV is HR-XML sine it has been being continuously developed. Nevertheless, its weakness point is that it has so many tiny details for CV in XML-based, such as birth place which it is not needed for the recruiters, and this point is corresponding to statements of CV guidelines advocating a good CV for job application (College of Law, 2010). Therefore, this study proposed the ontology corresponding to the Europortfolio Europass CV standard, draft version 2.0.3.0, which is widely used and more user-friendly uses to applicants. Fig. 1 shows the information of relevant standards used to represent a CV of an applicant. The details are described as the following.
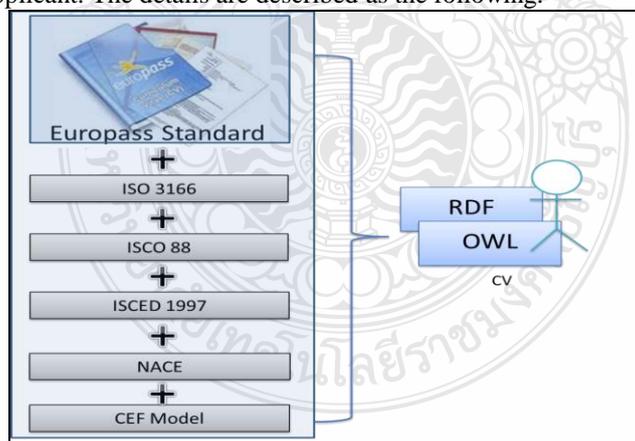

Fig. 1: The relevant standards used to represent a CV

- The ISCED 1997 is "International Standard Classification of Education", which represents education level (UNESCO, 2006).

[1] http://www.craigslist.org/about/sites/
[2] http://www.careerbuilder.com/
[3] http://www.indeed.com/
[4] http://www.monster.com/
[1] http://ramonantonio.net/doac/0.1/

- The ISCO 88 is International Standard classification of occupations (International Labour Organization, 2004), representing the desired employment and position in work experience of an applicant.
- The NACE code system is the European standard for industry classifications (Eurostat, 2008). It is used for representing the type of business where the applicant ever worked.
- CEF model is Common European Framework of Reference for Languages: Learning, Teaching, Assessment (Council of Europe, 2011). CEF separates the levels of language into six levels including A1, A2, B1, B2, C1 and C2; this starts from basic until proficiency respectively. It is used for representing the language skills.
- ISO 3166 is a standard published by the International Organization for Standardization. It is used for representing nationality of an applicant. Indeed, it will be useful in the system when users select a nationality from a list of "nationalities code"[2].

Besides, this research used various namespaces to describe the information resources, which are listed in the Table 1 below and then all of these namespaces are used during the implementation part. The information model was created based on OWL/RDF, which is a language for processing web information.

Table 1: List of namespaces

Namespace	URL
rdf	http://www.w3.org/1999/02/22-rdf-syntax-ns#
europass	http://openprofessional.ueuo.com/ontologies/europass.owl#
xsd	http://www.w3.org/2001/XMLSchema#
owl	http://www.w3.org/2002/07/owl#
nace	http://ec.europa.eu/eurostat/ramon/ontologies/nace.rdf#
rdfs	http://www.w3.org/2000/01/rdf-schema#
countries	http://www.daml.org/2001/09/countries/countries.daml#
isced	http://openprofessional.ueuo.com/ontologies/ISCED-1997.owl#
iso-3166-ont	http://www.daml.org/2001/09/countries/iso-3166-ont#
isco-88-com	http://openprofessional.ueuo.com/ontologies/ISCO-88-COM.owl#
foaf	http://xmlns.com/foaf/0.1/
nace_r2	http://ec.europa.eu/eurostat/ramon/rdfdata/nace_r2/
doac	http://kaste.lv/~captsolo/semweb/resume/cv.rdfs#
vivo	http://vivoweb.org/ontology/core#

Business Model

The research would like to propose the new business model of an online recruitment by applying the Semantic Web technology so that it can solve the mentioned problem in section 1. Fig. 2 shows the big picture of business model of the system, briefly explained as the following.

- Frist, applicants create their personal web site and then provide their profile (RDF CV) on the web. According to the benefits of RDF/OWL model, job-seekers can easily share their CV in various forms to others. Therefore, job-seekers do not need to fill their profile again because recruiters can find them as the concept of Linked Open Data.
- Second, applicants and recruiters register in the registration site, which is like an advertising site in order to propagate their requirements. Besides, applicants can make the white lists of recruiters, who are allowed to access their profile. For example, if an U1 (Applicant User 1) would like to participate with only three recruiters (R1, R2, and R3); U1 can make the white lists including R1, R2, and R3. Thus, R1, R2, and R3 can be permitted to access U1's information. On the other hand, other recruiters cannot access U1's information, to protect that information.

---

[2] http://www.iso.org/iso/iso-3166-1_decoding_table

- Third, recruiters can query their target candidates by specifying conditions in searching module. Indeed, the information of candidates is fetched from registration web site first and then it will be fetched from candidates' site so that it can increase the speed for querying.
- Fourth, the result of associated candidates will be shown.
- Finally, recruiters can go to the detail of their interesting candidates, and then decide later for choosing the target employee.
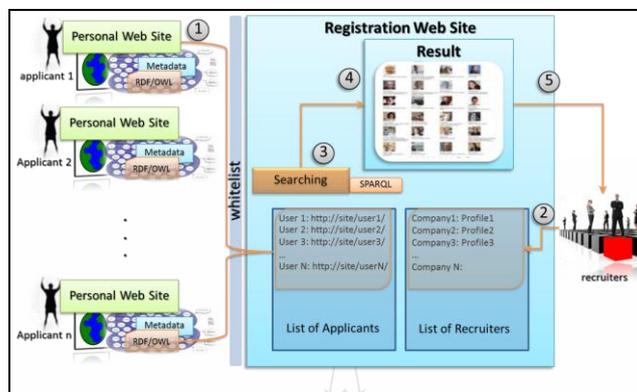

Fig. 2: The business model of the system

3. IMPLEMENTATION

In the system, applicants' information is stored in the form of RDF/OWL, which has more benefits to share. The proposed business model and information model in section 3 and 4 will be combined into the system, depicted in the Fig. 3, that show the overall architecture of the system by using Semantic Web technology based on the proposed business model.

The system is separated into the Open Professional site (registration site), which is like a medium system with the intention to retrieve and propagate RDF data (CV) of applicants and the personal site, which is only used to fulfil the requirements of the proposed business model that applicants can have their personal web site. The architecture of the prototype system is described as the following.

- Layer 1 is the user interface environment, providing the variously essential features.
- Layer 2 is the API modules based on Drupal[3], including RDF, CCK, EVOC, RDF CCK, ARC2, SPARQL query, RDF, and SPARQL endpoint module in personal site. In Open Professional site, it includes two APIs (ARC2 and RAP PHP). These are extensions connecting with the system so as to easily manage the metadata in the layer 3.
- Layer 3 is the metadata stored in the system for all relevant information in RDF/OWL form.
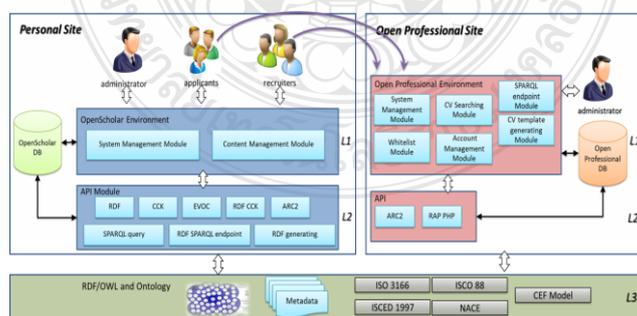

Fig. 3: The architecture of the prototype system

---

[3] http://drupal.org/

4. RESULTS AND EVALUATION

In comparison among exiting systems; the research proposed the criteria, which are stated below.

- *Creating individual site to organize his/her information and can convert to RDF CV.* Due to this research applying Semantic Web technology to enable the proposed business model.
- *Using ontologies and vocabularies to define and organize data.* As said by mentioned paragraph, lack of meaning of information is a weakness point in current recruitment system. For that reason, this research has tried to use benefits of ontologies for solving the problem.
- *Making white lists for individual target.* For the reason that data privacy is a serious problem in online system nowadays.
- *Centralized applicant's information.* RDF is an important resource for enabling data sharing among systems so that data is should not be duplicated.
- *Supporting general keyword for CVs or Jobs search.* This feature is commonly used for recruitment systems.
- *Supporting advanced CVs or Jobs search.* Likewise, this feature is commonly used for recruitment systems because it is an important feature in a regular recruitment system.
- *Supporting SPARQL query.* Due to SPARQL is a perfect query language for extracting information from RDF data.
- *Supporting Liked Open Data.* This criterion is also an important task, because the weakness point of most job recruitment systems is that they cannot share applicants' information or jobs' information together.

By the purpose in this study, only above criteria was used for comparison among the system with the popular job search engines according to suggestions of (List of Search Engines, 2012). The summary of comparison is shown in the Table 2.

Table 2: Summary of comparison among other recruitment systems

Criteria	Open Professional/ personal site (this research)	Monster	Job DB	LinkedIn	CareerBuilder	TaradJob
Creating *individual site* to organize his/her information and then *convert to RDF CV*.	√	×	×	0	×	×
Using ontologies and vocabularies to define and organize data.	√	√	×	×	×	×
Making white lists for individual target	√	1	1	√	1	1
Centralized applicant's information (data is not duplicated)	√	×	×	×	×	×
Supporting general keyword for CVs or Jobs search	√	√	√	√	√	√
Supporting advanced CVs or Jobs search	√	√	√	√	√	√
Semantic Search	√	√	×	×	×	×
SPARQL query	√	×	×	×	×	×
Linked Open Data	√	×	×	×	×	×

√ - can support
× - cannot support
0 - cannot generate RDF CV
1 - the CV can be only opened or closed

5. CONCLUSION

In sum, the result expresses that this proposed system is quite better than others, especially in supporting RDF, supporting ontologies and vocabularies, supporting SPARQL, and enabling linked data by SPARQL endpoint. These can solve the lack of a well-defined meaning of information and enable information-sharing across job portals. However, this research still has some limitations. For example, it does not have a model to represent recruiters or companies. Hence, following statements is explanation of suggested future works.

This study has designed only information model of CV for an applicant. Hence, the system has no advanced searching function to find a desired company. In the future, the system should include information model for representing a company, and then to provide an advanced searching function for finding a desired company. Also, in searching algorithm, it does not include the weight of factors in querying. Indeed, it should be included because there are many factors which can impact on choosing the best candidates depending on an aspect of each occupation. Therefore, if weight is given to each factor, it should be better. Maybe fuzzy-based is suitable to enhance for matching among applicants and recruiters.

References

[1]   Bizer, C., Heese, R., Mochol, M., Oldakowski, R., & Tolksdorf, R. (2005). The Impact of Semantic Web Technologies on Job Recruitment Processes. *International Conference Wirtschaftsinformatik.*

[2]   College of Law. (2010). *2010-2011 Resume Guidelines.* New York: Syracuse University College of Law.

[3]   Council of Europe. (2011). *Common European Framework of Reference for Languages*. Retrieved November 11, 2011, from Education and Languages: http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp

[4]   Dorn, J., & Naz, T. (2007). Integration of Job Portals by Meta-search. *International Conference on interoperability for Enterprise Software and Application.* Portugal.

[5]   Dorn, J., Naz, T., & Pichlmair, M. (2007). Ontology Development for Human Resource Management. *International Conference on Knowledge Management.* Vienna.

[6]   European Commission Group. (2010, December 3). *10 million use Europass CV for job-search.* Retrieved November 9, 2011, from European Commission: http://ec.europa.eu/education/news/news2692_en.htm

[7]   International Labour Organization. (2004, September 18). *International Standard Classification of Occupations.* Retrieved November 12, 2011, from International Labour Organization: http://www.ilo.org/public/english/bureau/stat/isco/isco88/index.htm

[8]   List of Search Engines. (2012). *List of the Best Job Search Engines.* Retrieved January 2012, 2012, from List of Search Engines: http://www.listofsearchengines.info/job-search-engines

[9]   Mirizzi, R., Noia, T. D., & Sciascio, E. D. (2009). A Semantic Web enabled System for Resume Composition and Publication. *IEEE International Conference on Semantic Computing*.

[10]   Robbins, S. P., & Coulter, M. K. (2009). *Management.* Prentice Hall.

[11]   Stevenson, B. (2008). The Internet and Job Search. *NBER Working Paper*.

[12]   Stone, J. R. (2006). *Human Resource Management* .

[13]   UNESCO. (2006). *International Standard Classification of Education.* UNESCO-UIS.