

Proposal for Job Portal Development

1st Duy Ta Bui
Faculty of Informatics
Pham Van Dong University
Quang Ngai city, Vietnam
(+84) 948 495 015
duybui.vn@gmail.com

2nd Long Phi Vu
Department of CSE
Maharishi University of Management
Fairfield, Iowa, United States
(+1) 720 840 2119
longvuphi88@gmail.com

ABSTRACT

Together with the development of internet, nowadays, websites of recruitment were become familiar with users. Job portal is considerably useful and more convenient for both job seekers and employers. However, the number of complaints users are leading to decrease the number of users. In this paper, we proposal the best technology in building for Job Portal which satisfied the requirements of users based on 7C Frameworks – Liferay Enterprise Portal. In addition, we would like to composite Recommender System and Matching System to connect user's requirement easier. The demo is called Job Matching Portal (JMP), will come up with a perfectly recruitment's website in the future.

Categories and Subject Descriptors

H.3.5 [Software and its engineering]: Organizations principles for web application; H5.2 [Information Interfaces and Presentation]: Users Interfaces---Evaluation/Methodology H.2.3 [Database Management]: Languages---Query; H.4.2 [Information Systems Applications]: Type of System---Decision support;

General Terms

Organizations principles for web application.

Keywords

Web portal, Liferay Enterprise Portal, 7C Frameworks, Dynamic Database Management System, Recommender System, Matching System.

1. INTRODUCTION

Each year, our country to receive the thousands of graduate's students and with them the first problem is looking for a suitable job. In addition, we have to mention tens of thousands of ordinary labors, technical workers.... Statistics also show that, from 2001 to now, the labor force growing at high speed; average increase 2.4% per year, equivalent to over one million workers.

There have been significant changes in awareness of job seekers in recent times. Besides knowledge and skills that workers have to equip themselves with knowledge about social law, the basic understanding of the field and career needs of labor market. Therefore, workers are already raising their autonomy in the process of looking for work, defines of "apply for a job" has replaced entirely by "looking for a job" which showed the initiative of workers in search, select suitable jobs with the ability and your needs. Recently, the establishment of large and small businesses with diversity recruitment demand has led to problems of finding employment is no longer a difficult matter for skilled

and qualified workers. In addition, the tendency to use the Internet is rapidly growing in the field of employment with more and more companies choose internet to advertise and recruit. According Richardson (2009), with the Internet, the recruitment process can be simplified the three processes as followed: faster job posting, faster applicant's response, and faster resume's processing.

As a result of above factors, it has led to the birth of the recruitment site. With useful and convenient features to use, the job recruitment site is being viewed as the optimal solution that workers and companies can meet their own demand. Currently, there are some big employment sites such as: "jobsite.co.uk" "indeed.com", ..etc. Although much work has been done to date, people still has problem with finding job because there are huge result of recruitment from these sites. More studies need to be conducted to minimize amount of results to help employers find their most suitable job. The critical requirements of the new model are: eliminating paperwork, improving time-to-hire, reducing turnover, creating a resume and position-centric environment as well as using the Internet as a recruitment and selection tool by matching features of system.

2. OVERVIEW TECHNOLOGIES

Based on the statistical analysis about 7C framework, in this session, we would to show a brief introduction about our technologies using in JMP:

2.1 7C Frameworks

Based on the complaints user received throughout 2013 of Indeed Inc, and Jobsite Inc, we would applying the research of Bens Paradamean with 7C Framework which was a guide for designing a commerce website (Rayport & Jaworski, 2008)[3].

In the research, a 35 question survey was created and grouped into eight sections: Context, Content, Community, Customization, Communication, Connection, Commerce and User's interest. The survey was distributed 125 job seekers and 125 employers. The questionnaire is measured by Likert scale; '1' for "Strongly Disagree", point '2' for "Disagree", point '3' for "Agree", and point '4' for "Strongly Agree".

The result data were analyzed using multiple linear regression technique. Chi-square analysis was also performed to further analyze which elements of the 7C framework were essential for the users. The results showed that only four out of seven elements from the framework affected the users' choice towards job portal, namely the context, the content, the community, and the commerce.

Table 1. Description of 7C Frameworks.

Element	Description
Context	The layout design, the look-and-feel of a user interface.
Content	The information provided by the website.
Community	How users of the website communicate to one another
Customization	The settings of a website. Background color, notifications, privacy etc.
Communication	How the website can facilitate interaction between the company and the user
Connection	How a website serves links to other websites to make it accessible by the users.
Commerce	How the website supports financial transactions between organizations and individuals

2.2 Liferay Enterprise Portal

Liferay Enterprise Portal is a free and open source enterprise portal written in Java and distributed under the GNU Lesser General Public License. It is primarily used to power corporate intranets and extranets. Liferay Portal allows users to set up features common to websites. It is fundamentally constructed of functional units called portlets. Liferay is sometimes described as a content management framework or a web application framework. Moreover, Liferay's support for plugins extends into multiple programming languages, including supporting for PHP or Ruby portlets. These applications are integrated together in a consistent and systematic way.

2.3 Why use Liferay Portal to build JMP

According to the result of 7C Framework, we chose Liferay Enterprise Portal to build JMP. In this part, we give 10 reasons why Liferay portal is the best product for JMP:

- Liferay has the lowest Total Cost of Ownership (TCO).
- Second-to-none rich out-of-the-box (OOTB) functionality around core portal, content management, collaboration, social, mobile, security and more..
- All portal products typically need extensions and/or additions to deliver requisite functionality – with Liferay you can simply do more within a specific budget.
- Product innovation – leader in introducing new capabilities whether it be AJAX or friendly URLs or mobile or social
- Improved business agility – it is lightweight in nature; you can quickly get it up and running, and it is easier to develop on/manage.
- A mature Enterprise Open Source (fully supported) product
- Liferay’s open architecture and its open source nature help you avoid lock-in to a single proprietary vendor.

- Liferay's hook and extension plugin model allows you to tailor product behavior to your needs without rewriting from scratch and without creating upgrade hell.
- Liferay offers opportunities for product feature sponsorship to enable contributions back into the core product for key customizations
- Liferay offers you a full choice of application servers, databases, and operating systems to run on, thereby allowing you to leverage your infrastructure and skills investment

2.4 Recommender System

Recommender engines (REs) also known as recommender systems are software tools and techniques providing suggestions to a user. The suggestions provided are aimed at supporting their users in various decision making processes such as what items to buy, what music to listen, what profiles to browse, or what news to read [3].

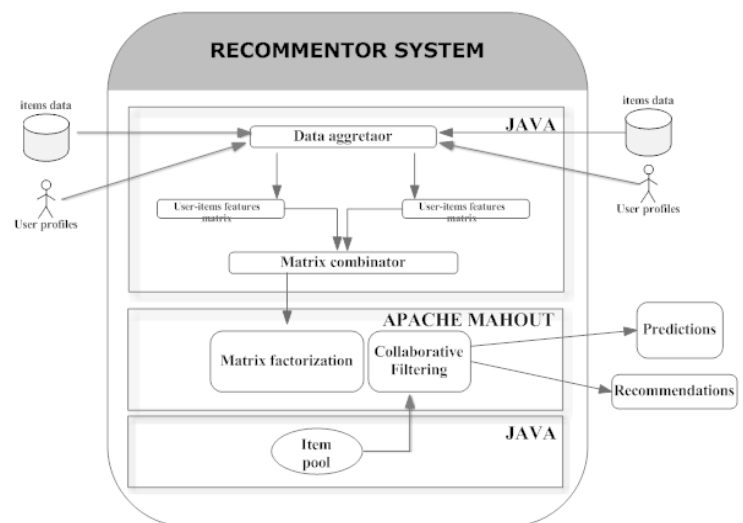


Figure 1. Overview of the flow of information throughout the Recommender System.

JMP applying the feasibility of the integration of a recommender engine as a module in a Liferay portal, and shows the process of its design and implementation using the Apache Mahout library. As such our work tackles two major problems which are: (1) the implementation of the recommender engine using the Apache Mahout library, and (2) the integration of the recommender in Liferay portal. The Figure 1 show an overview of the hybrid REs, and illustrates the flow of information through the system. According . F. Ricci and A, there are several REs, but we just suggested for 3 most popular taxonomies in JMP:

- Collaborative filtering: recommend to the active users with similar tastes liked in the past.
- Content-based filtering: the system learns to recommend items that are similar to the one that liked in the past. It is used two mains method to generate recommendations: Heuristic-based methods (TF-IDF) and model-based methods (use probabilities model)

- Knowledge based recommender system: Knowledge-based systems recommend items based on specific domain knowledge about how certain item features meet users' needs and preferences and, ultimately, how the item is useful for the user.

3. DYNAMIC DATABASE ARCHITECTURE

In JMP, we design database based on Dynamic Database Concepts to provide a certain degree of isolation between data structures and software that uses such as: Recommender System and Matching System. A dynamic database management system includes a data dictionary, a data importer and a query front-end. The data importer automatically imports data from an input file into a database, while adding new tables for new attributes as necessary, and updating parameters and folders tables in the data dictionary accordingly, so that end-users may access the imported data by database queries through the query front-end.[10]

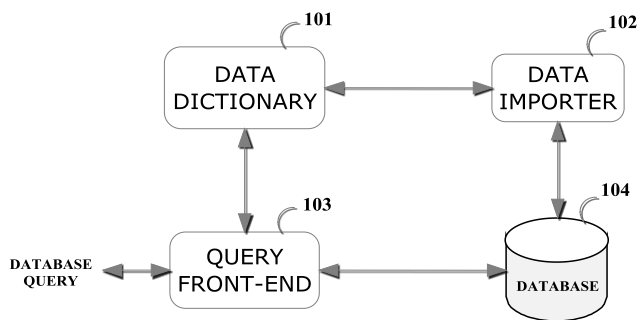


Figure 2. Diagram of a dynamic database management system.

To apply Dynamic Database concept, we use a items called Tag. Tag used to describe items. Tag is the data dictionary 101. Tag stores meta-data including table relationship and attribute locations. The data importer 102 writes data into the database 104 based on the information stored in Tag 101 and updates the data necessary after such writings. The query front-end 103 presents the database 104 to end-users in the form of parameter tree with users front-end address two issues. First, the end-user works with the names that are familiar to him or her, organized into custom folder structures adopted to the domain and/or usage of the data instead of dealing with the database directly, where the naming and organization of the data can be quite cryptic to the ordinary end-user. Second, the dynamically added parameters are immediately made available to the end-user and can be then moved to the appropriate folders by the database administrator or other authorized user.

And, the items of data importer will being depend on two actors:

- Jobseeker: provides information from users profiles. Jobseekers have the possibility to describe their own profiles information or job requirements by using tag.

- Employer: are institutions who publish or advertise job vacancies. They are provided information by search profiles or publish job vacancies. In addition, they also use tag to describe their own requirements.

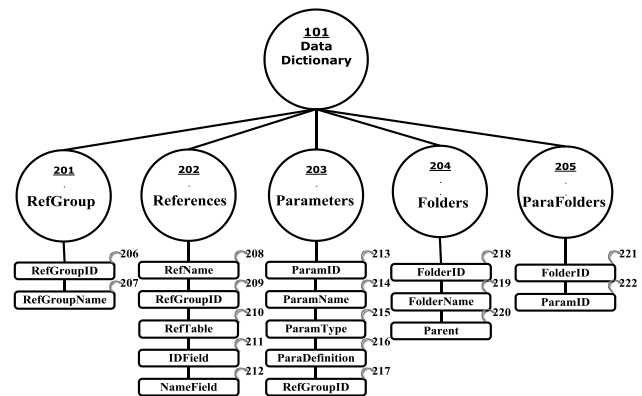


Figure 3. Diagram of dynamic management method.

Following this dynamic database structure, in JMP, we have database structure:

- RefGroup: RefGroupID is TagID. RefGroupName are table database in JMP such as: LearningOpportunity, JobVacancy, LWCArticle, Author, LFMArticle, LearningInstitution, Job, Language, Skill, Other(used to save special information without above items from users)
- References: RefName is new type of attributes which users want to add. IDField is attributes of tables in JMP. RefTable is TagName. NameField is the information which users want to provide.
- Parameters: are used to save information of new attributes. Folders: are used to save information of tables. ParaFolders: used to retrieval information of ParamID. The attributes of Parameter, Folder or ParaFolders are created automatic and hidden on user interface.

With this structure, users could describe new tag to provide their own information. The new information will not overlapping with database. In addition, it's open a new direction for providing by the documents. In this case, the data importers are the documents which are posted by users.

4. ADVANTAGES TECHNIQUES IN SEARCHING

In searching and ranking job, we use information retrieval (IR). The meaning of the term information retrieval can be very broad. However, as an academic field of study, IR might be defined that: "Information retrieval is finding material (usually document) of an unstructured (usually text) that satisfies an information need from within large collections (usually stored on computers)"[1]. IR has two basic models; there are set-theoretic models (typically standard Boolean model) and algebraic model or called statistical model (typically vector space model). In this section, we proposal steaming method by using Boolean model of IR to reduce the redundant information of documents.

4.1 Steaming process

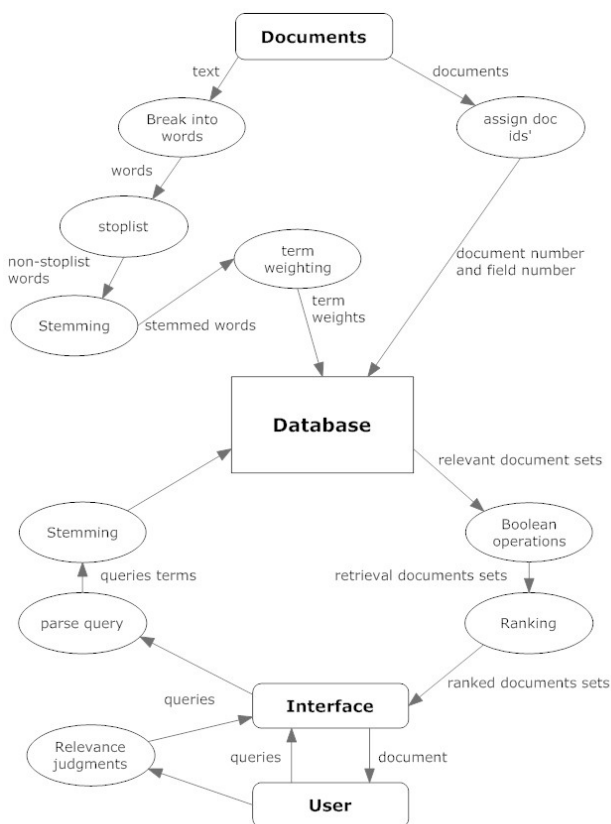


Figure 4. Steaming process for data redundancy.

Figure 4 describes the whole steaming words process in JMP. When building the database, job feature are taken one by one, and their text is broken into words. The words from jobs are compared against a stoplist (a list of words though to have no indexing value). Words from the document not found in the stoplist my next are stemmed. Words may then also counted, since the frequency of words in document and in the database as a whole are often used for ranking retrieved documents. Finally, the words are associated information such as the documents, fields within the documents, and counts are put into the database. The database then might consist of pairs of document identifiers and keywords as follows.

Keyword1 – document1-Field_2

Keyword2 – document1-Field_2, 5

Keyword3 – document3-Field_2, 2

Keyword4 – document3-Field_3, 4

.....

Keyword-n – document-n-Field_i, j.

Such a structure is called an inverted file. In an IR system, each document must have a unique identifier, and its fields, it field operations are supported, must have a unique field names. To search job, user enters a query consisting a set of keywords connected by Boolean operators (AND, OR, NOT). The query is

parsed into its constituent terms and Boolean operators. These terms are then looked up in the inverted file and the list of document identifiers corresponding to tem are combined according to the specified Boolean operator. If frequency information has been kept, the retrieved set may be ranked in order of probable relevance. The result of the search is then presented to the user.

We will explain and give examples about Boolean operations:

- **AND:** using AND narrows our search. It retrieves documents that contain both of the search terms or keywords that we specify. The more terms we connect with AND, the fewer search results we will find. For example, Software AND Engineer. The documents are retrieved containing both search terms. The shaded area represents search results.
- **OR:** using OR broadens our search. It retrieves documents that contain either of the search terms or keywords that we specify, but not necessarily both. The more terms we connect with OR, the more search results using it to search for similar terms. For example: Developer OR Liferay. Documents are retrieved containing either search term. Yellow shaded area represents search results.

4.2 MySQL full-text search cluster

MySQL has had FULLTEXT searching in one form or another since version 3.23.23. Full-text is built in functionality in MySQL that allows users to search through certain tables for matches to a string. Full-text indices in MySQL allow database administrators and programmers to designate any character-based field (CHAR, VARCHAR or TEXT) as a FULLTEXT index, which allows for complex text searching against data stored in those fields. [5].

In JMP, we use Boolean Fulltext Search – Fulltext search in MySQL Boolean Fulltext Search is performed for words. These query specifies what words must be present in every found row and what words must not. Complex expressions can be created with Boolean operators and parentheses. It means each row either satisfies and does it absolutely, results of comparison is TRUE – or does not satisfy it – and again, completely, value is FALSE. For each row Boolean Fulltext Search is easy to say whether it matches given search expression or not.

5. CONCLUSION

In this paper, we have proposed Job Matching Portal mean as the feasibility of the integration to composite both of Recommender System and Matching System as a module of Liferay portal. Following JMP, the module Recommender System was designed and implemented using Apache Mahout Library and the module Matching System was integrated by using MySQL full-text search cluster. Our focus in this paper was proposal a few newest technologies are being applied and developed in JMP. In the futures works, we would like to apply meta-data search algorithms based on MySQL full-text search cluster for Matching System. Especially, we will add more way to import data for users by using their own documents.

6. ACKNOWLEDGMENTS

This research paper is the development of thesis project. We would like to give a thankful to our supervisor Professor Phan Hoang and thank to Mr. Sudiana, Mr. Khoi Vo, Mr Nam Dau for conversation on the actual problems of recruitment website.

7. REFERENCES

- [1] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze. 2008. *Introduction to Information Retrieval*. Cambridge University Press.
- [2] Francesco Ricci, Lior Rokach, Bracha Shapira, Paul B. Kantor: *Recommender systems handbook*, Springer, 2011, pp 332 – 340.
DOI=<http://dx.doi.org/10.4018/ijwp.2014010105>
- [3] Sudiana, Ben Parademean, *Users' Interest Assessment on Job Portal*, International Journal Web Portal, Volume 6 Issue1, 01/2014, pp 65-75.
- [4] Richard Sezov Jr: *Liferay in action*, Manning, 2012, pp 63-209.
- [5] Poornachandra Sarang : *Practical LiferaY java-based portal application development*, Apress, 2009, pp 261.
- [6] Jonas X. Yuan: *Liferay portal systems development*, Packt, 2012, pp 7-8.
- [7] Connor McKay, Jorge Ferrer: *Liferay developer's guide*, LiferaY Inc. , 2011.
- [8] Duy Ta Bui, Long Phi Vu. *Job Matching System*, Thesis 04/2011.
DOI=<http://csc.hcmiu.edu.vn:8080/dspace/handle/123456789/541>.
- [9] Sean Owen, Robin Anil, Ted Dunning, Ellen Friedman: *Mahout in action*, Manning, 2012, pp 3-5-11
- [10] Hovhannes Ghukasyan, Suren Chilingaryan, Yervant D. Lepejian, *Dynamic database management system and method*, Granted Parent US 7174341 B2, 02/2007.
- [11] George Reese, Randy Jay Yarger, TimKing: *Managing and Using MySQL*, O'Reilly Media, 04/2002.