Job to Major (J2M): an Open Source Based Application

Haitao Liu

Department of Information Management
Dalian Neusoft University of Information
Dalian, China
E-mail: liuhaitao@neusoft.edu.cn

Jiacong Zhao*

Department of Information Management
Dalian Neusoft University of Information
Dalian, China
E-mail: zhaojiacong@neusoft.edu.cn
*The corresponding author

Dongzhao Zhou

Department of Information Management
Dalian Neusoft University of Information
Dalian, China
E-mail: zhoudongzhao@neusoft.edu.cn

Shuang Gao

Department of Information Management
Dalian Neusoft University of Information
Dalian, China
E-mail: gaoshuang@neusoft.edu.cn

Abstract—**This paper presents Job to Major (J2M), an open-source tool design to link job requirements with University of Southampton's major information. J2M currently provides two functions: suggesting the most suitable majors to potential students of the university based on the job requirements, find jobs that are closely link to what students acquire from their major. The development of J2M is according to Garrett's model which divides the application development mainly into three stages business goals analysis, system design and implementation. J2M build on existing open standards and supported by open-source of University of Southampton, Universal JobMatch provided by gov.uk and additional tools. The evaluation of J2M tries to show the functions achievement of J2M.**

Keywords- Job to Major (J2M); Open innovation; Business; Design; Implementation

I. INTRODUCTION

Currently, both students and universities are placed in dilemma in terms of the relatedness of college major and job market. Firstly, students attend college and select degree fields in the hope of succeeding in the labor market [1]. Furthermore, the ability to utilize the investment in schooling in future employment [1,2,3] is one aspect of labor market success. Hence, most potential students chose their majors based on requirements of the job. However, it is difficult for them to compare which majors of the university contribute most to their future career through general information searching. In addition, occupation specific skills related to the current occupation increase wages [2]. In contrast, wages are lower for mismatched workers who are working in a job that uses fewer of the occupation specific skills learned by graduates in the major. This dilemma is described as mismatch between education and job market [1]. Thus, in order to make what they have learnt valuable and get relative higher salary, graduates try to find a job that highly corresponds to what they have learned. Secondly, it is unreality to require universities to keep their majors' content up to date with the rapidly changing job market requirement based on traditional access to job information. Current solutions to students' career development can mainly be divided into three types. One is web applications like Total-Jobs and Target-Jobs are only focus on job finding. One is application like C4S, it only pays attention to major finding and the application is not user-friendly. The third way is the employment center of university which is both time and cost wasting. Thus, in order to address these problems, a new application should be proposed to link job requirements with students major information.

Open innovation allows organizations to look beyond their internal resources to develop new products, services, and sources of revenue, so as to solve real world problems [3]. Additionally, open data can be accessed and used freely by anyone and open data innovations usually contribute more opportunities to digitize, automate and optimize many kinds of current services[4]. In order to benefit to students, universities, and even parents and employers, the proposed application will based on the open source of job requirements and college majors' information. Therefore, we propose a open-source web application Job to Major (J2M) to solve the mentioned dilemma for students of University of Southampton. Open data innovation project delivered a rich and varied program of activities, like business, technology and even ethics. Thus, the development of J2M was organized into five sections. The following sections firstly demonstrate the theoretical background of open data innovation. Secondly, specifying the design. Then, turning to the implementation of J2M and following the texting of the whole system. Finally, evaluating the application through business feasibility, social contribution and techniques aspects.

II. THEORETICAL BACKGROUND

As the development of information technology, every company is exposed to a widely distributed knowledge world. Companies cannot afford to rely entirely on their own

research, in this situation, many innovative firms have shifted to an 'open innovation' model, using a wide range of external actors and sources to help them achieve and sustain innovation [4]. In this way, companies can reduce cost for conducting research and development, improve development productivity, increase the accuracy for customer targeting [5]. Additionally, the external resources generally gain by collaborating with local governments, universities, business support services, and other public bodies [6]. It indicates that open innovation businesses often shoulder social responsibilities. However, there are still existing kinds of problems in this process. For example, revealing information not intended for sharing and revealing intellectual property of hosting organization. In order to reduce these risks, how to exploit openness for firms' benefits has been the heart of recent research on innovation [4, 7, 8]. In this situation, open innovation models are proposed to formal the whole development process which includes idea generation, business analysis, application design, application implementation, testing and evaluation. For example, Garrett's model [9] focuses on user experience achievement. It divides the innovation process into five stages: surface, skeleton, structure, scope and strategy. These stages clearly define what goals the innovation try to gain, what tasks should do to support goals and what tools should be used to put tasks into practice. Fugle model [5] divides this process into seven stages, compared with Garrett's model, Fugle focuses more on concept establish and specifics platforming criteria for technical level [10]. It is therefore can be seen that current open innovation model can support the open-based web application development. J2M is developed based on Garrett's model and also considers Fugle model's platforming criteria.

The development of Web innovation applications should also be supported by robust web techniques. Majority of people spend a lot of time on Web. This motivates companies take diverse number of actions for open innovation [11]. For example, Procter & Gamble developed its "Connect + Develop website" to get in touch with external innovators, so that they could contribute through propositions of solutions to P&G problems. OpenCalais is also a kind of open Web service to support text analysis for J2M. With the development of web tools, Web has become a platform for collaboration which triggered the growth of open innovation platforms. These platforms try to leverage the Web technology and most notably its social aspects to help web innovation [11], like Node.JS and Express.JS. Current research also addresses paradigms in open innovation processes, on which we will base our further analysis (Table.1). These paradigms are applied to different research situations to deal with unexplored aspect and consequences in the innovation process. Furthermore, there are many identified Web technologies that are likely to be useful in problem solving processes on the open innovation platforms, like expert finding [12], semantic keyword matching [13, 14] and social propagation [15]. J2M mainly relies on the semantic keyword matching to broad the space of matching possibilities between major information and job requirements. Finally, tools like Sublime Text, SQLite

Manager and CSV to JSON Converter are mature enough to deal with objectives like code edit, data storage and data format transformation to support the goals achievement. Previous theoretical background analysis shows that current open innovation model and related web techniques guarantee the development of J2M is practical rather ambitious.

III. DESIGN OF J2M

In order to guarantee the productive development process of a project, it is better to make a plan. Gantt Chart Project Plan will be given firstly. In order to gain high user experience, Garrett's model is introduced for designing J2M, and J2M will be developed through the following three aspects: business strategy, system design and implementation.

TABLE I.	TABLE I. OPEN INNOVATION PARADIGMS
Social Behavior on the Web-"Weak Ties" [16]	Explaining the social behavior of people on the Web and to the Open Innovation process.
Cross-Sectorial Problem Solving [17]	Creating solutions to the problem that were at the border or outside companies' area of expertise.
Broadcasting [17]	Representing the distribution of different content related to problems to a dispersed audience

A. Business Analysis

The business goals of J2M is firstly to attract potential students of University of Southampton to use this application to find the suitable major based on the requirements of jobs that they are interested in. Then, attracting current Southampton students to use this application to find a job that highly consistent to what they have learned. In order to guarantee that J2M is customer focused and can survive in the market. Two business models VRINE (Value, Rare, Inimitable& Non-substitutable, Exploitable)[18] and SWOT (Strength, Weakness, Opportunities, Threats) [19] are introduced to analysis the resources and capabilities of J2M. The following five components (Fig .1) are identified important to the sustainable development of this application.
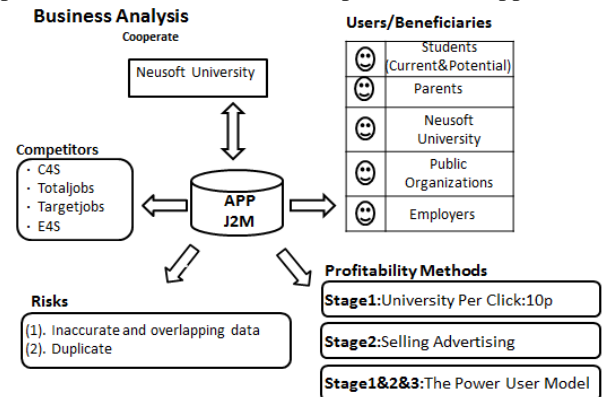

Figure 1. Business Analysis of J2M

1) Target Organization

A starting point for the idea of openness is that a single organization cannot innovate in isolation. It has to engage with different types of partners to acquire ideas and resources from the external environment to stay abreast of competition (Fig .2 [20]) [4, 8]. So we cooperate with Southampton University and develop this application to assist the university in terms of students' career development.
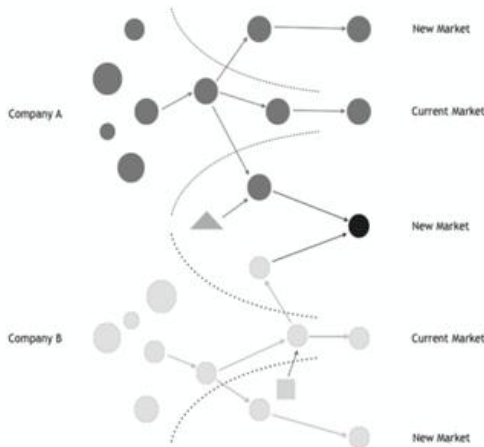

Figure 2. The Open Innovation Process [20]

2) Customers

Customer refers to the benefiter of our application. Potential students, based on their dream jobs' requirements, we suggest suitable majors of Southampton University to them. Current students can find jobs that highly match to the acquirements of their majors. Education industry has developed more similar to consumer goods market [21]. University provides its service and reputation to students. This market is also seriously competitive. In order to achievement sustainable development in this market, the university should sensitive to job market. J2M enable the university gain this advantages. Public organizations, like the education department with our data can track the current situation between education and job market. This helps them plan the future education project more practically. Employers usually spend a great deal of time and money to hunt for desired employee. J2M narrows the hunting scope for them, and they just need to pay more attention to specific major students. This is both money and time saving.

3) Competitors

Recently, four competitors are identified. Firstly, C4S is limited in major searching and has no fitting function. It is not user-friendly. The Total-Jobs, Target-Jobs and E4S are only for job searching. It can be seen that they all focus on major or job searching but not link them together. Our application fills this gap, which distinguishes us from these competitors.

4) Risk

The risk of J2M mainly consists of two parts. One is the dataset itself. The open dataset contains inaccurate and overlapping data, which will easily lead to inaccurate and even wrong output. This risk can be avoided by data refine tools, such as Google Refine. The other kind of risk is mainly from the duplication. Because of the easily access to open source, every can duplicate the proposed idea.

5) Profitability method

We design our profitability strategies into three stages:

- **Stage1:** Charge Southampton University by each click.

- **Stage 2:** When we gain relative more users then we can sell some advertising space to the website like C4S and total-jobs.

- **Stage 3:** We firstly provide good but limited information to the users in the previous two stages. Then charge users with the premium service.

B. J2M Systme Design

J2M owns two functions: returning a list of possible majors when a job title is input (Fig .3) and returning a list of job titles with description when a major title is selected (Fig .4). To keep implementation works in an appropriate scale, the study program should be limited to Electricity and Computer Science courses in the University of Southampton only. In the same reason, the kinds of job are also filtered to computer and IT related jobs. The detailed structure is as follows.

In Fig .3, when a job title is input to the interface, then it connects to open data Universal JobMatch provided by gov.uk. to filter related jobs, with short job descriptions. Then those job data will be connected to OpenCalaris, the open keyword finder to get keywords for each job. There will be another database in the application, the list of the keywords for each module in the University of Southampton. Those will be already listed up. Then the application will compare the keywords list of each job to the keywords of each module. Finally the list of possible study programs suitable to the job title will be displayed in the interface, with number of hit of keywords.
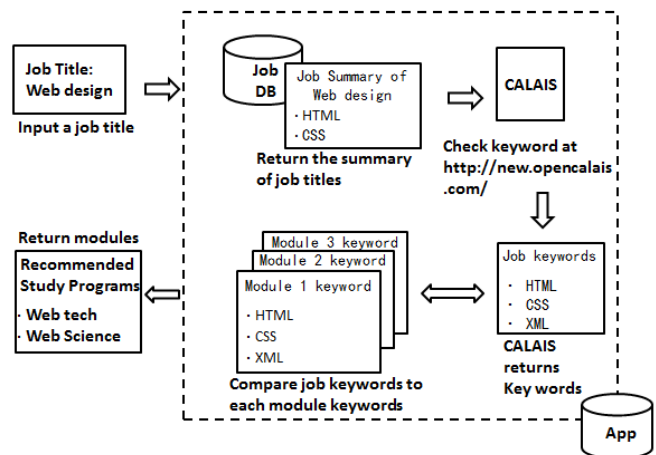

Figure 3. Major Sele cting

In Fig .4, The application works inversely. There will be a dropdown menu of study programs of Electricity and Computer Science in the University of Southampton. When a major is selected, the application will connect to the list of keywords for each module to get temporary keywords. Then it connects to the job database and OpenCalarsi to get keywords match to the temporary keywords, and returns the list of recommended job titles.
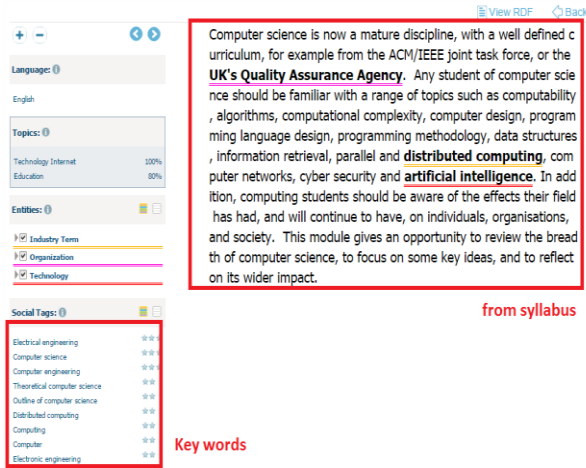

Figure 4.   Job Selecting

IV.   J2M IMPLEMENTATION

The application is modularized. There are modules for connecting to text analysis service, connecting to database to get course module keywords and matching job keywords and course module keywords. Ajax is used for gaining information or refreshing data on the webpage. This can avoid being disturbed by refreshing page and make manipulation more quick and smart. The following part mainly describes the function from job information to module. And the function from module to job is the same.

A. Frameworks and Development tools

The frameworks and development tools are the foundation for the application development. To J2M, the frameworks for developing are Node.js and Express.js. Development tools are shown in Table 2.

TABLE II.   TABLE II EVELOPMENT TOOLS

Development tools	Function
Sublime text	Code editing
SQLite manager	Data management

CSV to JSON Converter	Convert the open data format into .JSON
Sqlite3	Database for storing keywords of course

B. Open Data for Job

Job list is gained from Universal JobMatch website by using API (api.lmiforall.org.uk/api/v1/vacancies/search). Keyword user inputs will be sent through API to search jobs. After gaining job information from API, job titles were picked from the JSON data, and shown in table. Job information of each job is allocated using jQuery click event function. Job description of a specific job can be achieved by clicking job titles.

C. Open Data for Major

Firstly visit Website http://data.southampton.ac.uk/ dataset/courses.html for major information and download courses.csv file. Then, select module information by filtering out "DEPT_DESC is ECS" and "SUBJ_CODE is COMP" then extract unique PROG_CODE, to get 30 majors related to computer science. For each major, visit university site of module list, for example to get syllabus of all possible modules by http://www.ecs.soton.ac.uk/ programmes/msc_computer_science#modules.Fig .5 shows a model example.
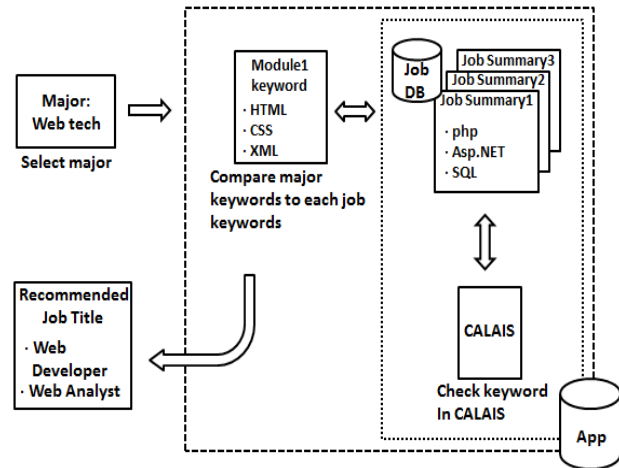

Figure 5.   An example of OpenCalais text analysis

D. Gaining Keywords

Job keywords for matching with module keywords are achieved by sending job description to OpenCalais text analysis service. Job description is a part of job information gained from API. OpenCalais API is connected by using Calais node module. Calais node module sends text to analyze and fetches text analysis result. Text is sent by HTTP request using jQuery Ajax. Course module keywords are gained from database. Fig .6 shows an example of OpenCalais text analysis.With OpenCalais the keywords for each model can be achieved, and then list up all keywords for each major (e.g. Fig .7). Then, list up analyzed keywords with major(module) code (e.g. Fig .8).

ID	keyword	course
00001	3-D	4431
00006	access Wireless sensor networks	4431
00009	Actors Timing Hardware	4431
00014	Africa	4431
00016	AJAX	4431
00019	AMAZON, INC.	4431
00023	animation	4431
00027	artificial intelligence	4431
00028	ASP.NET	4431
00029	ASP.NET Technologies	4431
00039	basic utilities	4431
00041	Bayesian Neural Networks	4431
00042	beam search	4431
00043	belief systems	4431
00046	Bioinformatics	4431
00050	Biosensors	4431
00051	BITTORRENT INC	4431
00055	Bootloader Stack	4431

Figure 6.   An example of module keywords

4431	4432	4433	4434	4435
3-D	3-D	3-D	3-D	3-D
access Wi	access Wi	access Wi	access Wi	access Wi
Actors Tim	AJAX	AJAX	AJAX	Active an
Africa	AMAZON,	AMAZON,	AMAZON,	Actor Net
AJAX	animation	animation	animation	Adam
AMAZON,	ASP.NET	ASP.NET	ASP.NET	adaptatio
animation	ASP.NET T	ASP.NET T	ASP.NET T	Africa
artificial i	Bayesian I	Bayesian I	Bayesian I	agent-bas
ASP.NET	Bioinform	Bioinform	Bioinform	AJAX
ASP.NET T	biometric	biometric	business t	AMAZON,
basic utili	business t	business t	C++	AMDAHL
Bayesian I	C++	C++	CGI	animation
beam sea	CGI	CGI	client-side	artificial i
belief syst	client-side	client-side	client-side	ASP.NET

Figure 7.   Keywords with major (module) code
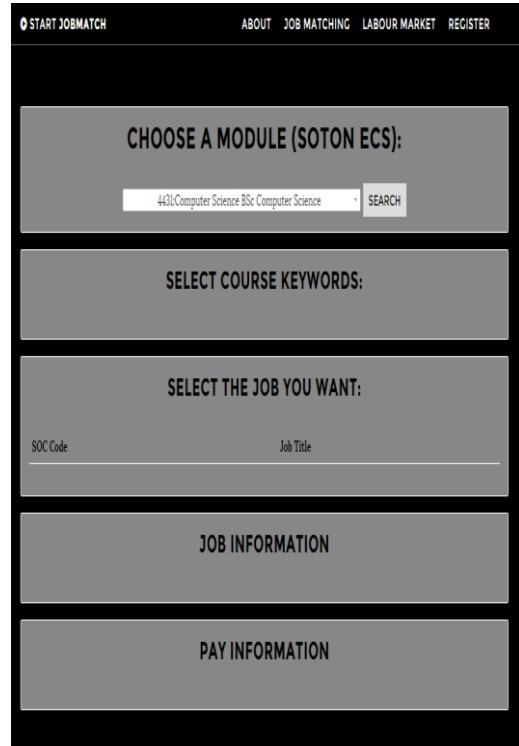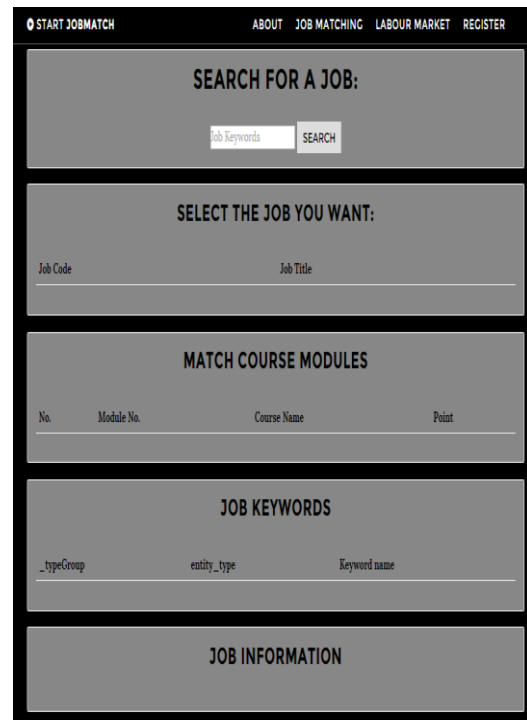

Figure 8.   Figure 8. Job Working

Figure 9.   Job Finding

E. Getting Module Recommendation

Module recommendation is made based on the result of keyword matching. Job keywords are sent to matching function module and matched with module keywords. Matching point will be added once keywords match. Matching point calculated in this function will be returned and sorted to recommend the most matched modules.

The testing here is to test whether J2B can achieve functions for job and major selecting.

1) From job requirements to major

When users search job keywords, the "SELECT THE JOB YOU WANT:" part will show you the jobs. After users selecting a job, the rest parts will show the matched courses with points, job keywords and job information.

2) From major to job position

This part is to test whether J2M can find a job position based on the major information (Fig . 9).

The test result shows that J2M can achieve these tow functions well.

V. J2M EVALUATION

The evaluation of J2M will be discussed through two aspects. One is for the evaluation for a web application which should achieve high level of design schema, information content and ease-of-use. The other aspect is for open innovation evaluation. This refers to compared with traditional close data application J2M brings what kinds of discovery and divergence.

A. Web Application Evaluation

1) Design schema analysis

This part only focuses on static descriptions of the application, which mainly verifies the correctness and consistency of system design specifications [22,23], to enhance the quality of conceptual schemas by looking for design inconstancies and irregularities in the application of design patterns [24]. The patterns addressed by J2M consist of compositions of hypertext elements: pages, units, operations and links. These elements are typically serving application purpose. The testing part shows that J2M can arrange of pages, units, and links for supporting the navigation between job and major information. And in J2M a core object can be accessed via one or more access objects. For example, a job position can be directly accessed via module content key words or from module number which needs more accesses. Hence, J2M achieves acceptable design schema.

2) Information content evaluation

J2M's information will be evaluated by criteria provided by [24]. Five types of criteria are discussed: orientation information to Website, Content Information, Metadata, Services, Accuracy. Table.3 clearly shows that J2M achieves most of the criteria in the information contents.

TABLE III. INFORMATION CONTENT EVALUATION CRITERIA AND J2M ACHIEVEMENT

Evaluation Criteria for Information Content	Achievement of J2M
Orientation Information to Website	
A website overview is provided: States purpose/mission of website, appropriate to entity's overall mission	YES
Scope of website is clearly stated: Type and origin of information, audience, dates of coverage, etc.	ALMOST YES
Services and information provided at the website are described.	ALMOST YES
"What's new" section: alerts frequent users to changes in content, services, etc.	NO
Instructions for the use of the website are provided	NO
A liability/status statement warning the user of the nature of information provided at the site, and through any links made from the site, is provided	YES
Copyright statements are provided	YES
Content Information of Website	
Match the purpose/mission	YES
Match needs of stated audience	YES
Includes only necessary and useful information	YES
Coverage does not overlap: within the site, or with other agencies	YES
Amount of information is significant, and balanced.	YES
Contains direct information resources	YES
Clear and consistent language style that matches audience: Plain English, use of Maori, Pacific islands and Asian languages if appropriate	NO
Positive professional tone: Avoids jargon, inappropriate humour, condescension, accusation and chit chat.	YES
Content does not show bias: Racial, cultural, political, commercial	YES
External links are to appropriate resources, connected with the business of the entity	YES
Metadata: Facilitates retrieval, navigation	
Appropriate metatags are provided, e.g. title, author, description, keywords	YES
Headings are clearly phrased, descriptive, and understandable	YES
Each page is titled clearly	YES
Terminology and layout are consistent within the headings throughout the website	YES
Services	
Availability of services: open to everyone on Internet, or require fees, restricted to particular sector groups	YES
Meet needs of user	YES
Fully operational	YES
Accuracy	
Information provided is accurate	YES
Statement of status of document/website provided	YES
Sources of information are cited (accurately)	YES
Typing, spelling, grammar, and consistency errors are absent.	ALMOST YES

3) Web usage analysis

Web usage analysis refers to analyze dynamic data that collected at runtime and produce quality reports on content access and navigation sequences [23]. To J2M usage analysis, links, feedback, accessibility and navigability based on the criteria provided by [22, 23, 24] will be evaluated It can be seen that although J2M can support the job or major searching. But it should be update to a more user-friendly application.

B. Innovation Evaluation

The innovation of J2M will be evaluated based on the two fields ground by Garrett's model: business goals and technology. Firstly, traditional close data based business usually benefit to the business conducting company by exploiting customers. Compared with based business, the goals of open business is not only benefit to business conducting companies, but more customer and social responsibility focused. This means that open business tries to establish a win-win commercial environment, where every participant is beneficiary. For example, J2M benefits to a wide group of people as well as itself. Secondly, in terms of technology, one of the significant contributions is concurrency. Unlike closed data application, J2M does not

Need to update its data set frequently. Because, the dataset of J2M is concurrently with job and major information. However, disadvantages exit as well. For example, open innovation should choose the open source seriously, to avoid the inaccurate and fault outcomes caused by the original dataset.

The previous analysis shows that in technology field J2M achieves most of the evaluation criteria. What J2M should improve is to be more user-friendly. As an open data application J2M enjoys both business and technology advantages. However, J2M still needs to pay attention to risks for open data application.

REFERENCES

[1] J. Robst, "Education and job match: The relatedness of college major and work,"Economics of Education Review, vol. 26, 2007, p. 397407.

[2] Bauer, T. K., Educational mismatch and wages: A panel analysis. Economics of Education Review, 21(3), 221–229, 2002

[3] Shaw, Kathryn L. "A formulation of the earnings function using the concept of occupational investment."Journal of Human Resources (1984): 319-340.

[4] Chesbrough, H. W. (2003). Open innovation: The new imperative for creating and profiting from technology. Harvard Business Press.

[5] Du Preez, N; Louw, L. 2008. A Framework for Managing the Innovation Process, Picmet 2008.

[6] Enkel, Ellen, Oliver Gassmann, and Henry Chesbrough. "Open R&D and open innovation: exploring the phenomenon." R&d Management 39.4 (2009): 311-316.

[7] Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. Strategic management journal, 27(2), 131-150.

[8] Dahlander, L., & Gann, D. M. (2010). How open is innovation?. Research policy, 39(6), 699-709.

[9] Garrett, Jesse James. "The Elements of User Experience." Jjg. Net (2004).

[10] Marais, S. J., and C. S. L. Schutte. "The development of open innovation models to assist the innovation process." 23rd Annual SAIIE Conference Conference Proceedings. 2009.

[11] Jesic, D., Kovacevic, J., & Stankovic, M. (2011, June). Web technologies for open innovation. In Proceedings of the 3rd International Web Science Conference (p. 20). ACM.

[12] Aleman-Meza, B., Bojars, U., Boley, H., Breslin, J. G., Mochol, M., Polleres, A., et al. 2007. Combining RDF vocabularies for expert finding. Lecture Notes in Computer Science, 4519, 235. Springer. Retrieved from http://www.springerlink.com/index/p6u10781711xp102.pdf.

[13] Ziegler, C.-N., Simon, K., and Lausen, G. 2006. Automatic Computation of Semantic Proximity Using Taxonomic Knowledge Categories and Subject Descriptors. CIKM '06 Proceedings of the 15th ACM international conference on Information and knowledge management (pp. 465-474). Arlington, Virginia, USA:ACM New York, NY, USA. Retrieved from http://doi.acm.org/10.1145/1183614.1183682.

[14] Cilibrasi, R. L., and Vitanyi, P. M. B. 2007. The Google Similarity Distance. IEEE Transactions on Knowledge and Data Engineering, 19(3), 370-383. doi: 10.1109/TKDE.2007.48.

[15] Hastings, G. 2007. Social Marketing - Why should the devil have all the best tunes? Elsevier Ltd.

[16] Granovetter, M. 1983. The strength of weak ties: a network theory revisited. Sociological Theory, Volume 1 (1983), 201- 233.

[17] Aleman-Meza, B., Bojars, U., Boley, H., Breslin, J. G., Mochol, M., Polleres, A., et al. 2007. Combining RDF vocabularies for expert finding. Lecture Notes in Computer Science, 4519, 235. Springer. Retrieved from http://www.springerlink.com/index/p6u10781711xp102.pdf.

[18] Carmeli, Abraham. "Assessing core intangible resources." European Management Journal 22.1 (2004): 110-122.

[19] Hill, Terry, and Roy Westbrook. "SWOT analysis: it's time for a product recall."Long range planning 30.1 (1997): 46-52.

[20] Laursen, K., & Salter, A. (2006). Open for innovation: the role of openness in explaining innovation performance among UK manufacturing firms. Strategic management journal, 27(2), 131-150.

[21] Singh, M., "E-Services and Their Role in B2C E-Commerce', Journal of Managing Service Quality, 12:2, pp.434 – 446, 2002.

[22] S. Comai, M. Matera, and A. Maurino. "A Model and an XSL Framework for Analysing the Quality of WebML Conceptual Schemas". Proc. of IWCMQ'02 - ER'02 International Workshop on Conceptual Modeling Quality, Tampere, Finland, October 2002.

[23] P. Fraternali, M. Matera, and A. Maurino. "WQA: an XSL Framework for Analyzing the Quality of Web Applications". Proc. of IWWOST'02 – ECOOP'02 International Workshop on Web Oriented Software Technologies, Malaga, Spain, June 2002.

[24] Fraternali, Piero, Maristella Matera, and Andrea Maurino. "Conceptual-level log analysis for the evaluation of web application quality." Web Congress, 2003. Proceedings. First Latin American. IEEE, 2