



Ontology based multi criteria recommended system to guide internship assignment process

Abir M'Baya, Jannik Laval, Néjib Moalla, Yacine Ouzrout, Abdelaziz A Bouras

► To cite this version:

Abir M'Baya, Jannik Laval, Néjib Moalla, Yacine Ouzrout, Abdelaziz A Bouras. Ontology based multi criteria recommended system to guide internship assignment process. 3rd IEEE International Conference on Computational Science and Computational Intelligence (CSCI 2016), Dec 2016, Las Vegas, United States. hal-01788507

HAL Id: hal-01788507

<https://hal.science/hal-01788507>

Submitted on 6 Nov 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Ontology based multi criteria recommended system to guide internship assignment process

Abir M'baya¹, Jannik Laval², Nejib Moalla², Yacine Ouzrout², Abdelaziz Bouras³

^{1,2}Université de Lyon, Université Lumière Lyon 2, DISP Lab EA4570
Lyon, France

³CSE Department, Qatar University, Doha, Qatar

²FirstName.LastName@univ-lyon2.fr, ²FirstName.LastName@qu.edu.qa, ¹A.Mbaya@univ-lyon2.fr

Abstract—Internship component is a vital part of the university training program as it looked as a vital resource for students to gain the required skills for employment. Recognizing the importance of drawing a compromise between companies' requirements and students' skills in the internship assignment process has triggered the need of a multi criteria decision system that enables to select the right student to the right internship. Both competencies required by the recruiters and the students' skills evolve in the time and their regular update is necessary. This leads to the development of a reference model for their management, update and maintenance. This paper proposes a multi criteria decision system which aims at integrating an ontology in order to select and to recommend adapted internship seekers to the recruiter mission posting or vice versa. This proposed recommended system has four phases in screening candidates for recruitment. In the first phase, mission requirements are represented as ontology. The system collects candidates' resumes and constructs ontology model for the features of the students in the second phase. In the third phase, we construct the training courses ontology. Finally, we discuss the steps of the ontology based multi criteria decision making method that enables to retrieve the right candidates to the right internship.

Keywords- student profile; internship posting; semantic matching; ontology; screening candidate.

I. INTRODUCTION

Companies and more generally the economic world require from the school and universities the development of training systems more in line with their concrete needs in term of human resource. An interesting way is to establish an educational training system that enables to combine classroom based-education with practical work experience and help young people to make the school-to-work transition.

In this field, many universities provide an alternative education or "co-operative" courses or a work-link training in which the students can develop skills learned in the training academic program, expand their knowledge through related work experience, and explore career options and network with potential employers.

University Lumière of Lyon (ULL) follows this original mode of training such it helps students to develop their potentials through their practice in companies. In such a way, each student must have its appropriate mission in order to complete his training and validate his diploma.

Actually, universities often receive resumes from candidates and manually short list qualified applicants that

correspond to the missions' requirements. They therefore check the compatibility of the training programs with the internship postings that are written in a free text and then they identify and select students who match the qualifications required to perform a defined mission in the best way.

Internship assignment is a multi-criteria decision-making problem that includes both tangible and intangible factors. These problems refer to how structure all the information relating to internship offer and students' qualifications in a standardized structure that allows the migration, reuse and automatic analyze of the documents and then how students are assigned to a predefined mission in order to establish a balanced matching between business needs and students' capacities.

In order to alleviate these issues, we propose in this paper an ontology based multi criteria decision making recommended system that enables one to define decision models using ontology as the base construct.

A Recommended System is a tool aims at providing universities with useful information results searched and recovered according to their needs, making predictions about matching students to their appropriate missions and delivering those items that could be closer than expected. To do this, the student's information profile and companies' needs in term of human resource require to be stored. Ontology based multi-criteria decision making recommended system is a qualitative decision making system that structures decision models in such a way that the problem solution can be obtained by reasoning upon the ontology. The internship assignment ontology represents a modeled and structured knowledge associated with business requirements defined in missions and student profiles using semantic machine-interpretable concepts and relationships in such a way that they can be used by machines not just for display purposes, but also for processing, automation, integration, and reuse across system. OWL is used as the ontology representation language. Besides the OWL language, SWRL rules can be used for additional expressivity. Any available reasoning engine which supports the OWL language and SWRL rules can be used for reasoning support. Because concepts of OWL ontologies and concepts of multi-criteria decision making do not have a direct translation, a requirement for defining the ontology based decision making process was unambiguous correspondence of decision model elements in ontology-based and multi-criteria decision making system.

The aim of this paper is to provide an ontology based recommended system that brings semantic to the ordinary recommended system with formalized knowledge and data. It represents a great opportunity for universities that allows them to deal with these modeled data in order to recommend the best mapped student resumes suitable for the internship requirement using similarity measurement. i.e. this system helps companies in recruiting top talents and students in selecting the appropriate internship which corresponds to their skills and interests and building more easily their professional project. The main advantages of our ontology based recommended system are: a higher level of decision support provided by the ontology, a higher level of business process automation, direct use of information captured in ontology for decision making and reuse of decision model concepts in definitions of more complex ontology concepts.

The rest of the paper is organized as follows. In the next section, background information and related work are presented. In section II, we outline the ontology multi criteria based recommended system in which we define the domain ontology of three main parts: Profile of the student, training education and internship posting and present the different steps of the ontology based multi criteria decision making recommended system. Section IV describes the implementation overview while section V concludes the paper and suggests future works.

I. BACKGROUND

Domain ontology emerged in mainstream in many applications. Recently it has increased the importance of applying ontology as a key part of efficient filtering in recommendation systems [5], [6]. In this section we define the most important terminologies used in this paper. Then we review the related works in e-recruitment and ontology mapping that is considered as the most relevant to our research. However, there has been relatively little research exploring semantic balanced matching systems between employers' and candidates' preferences. Besides, ontology-based decision making has not yet been discussed from our perspective.

A. Ontologies

Ontology is an explicit formal specification of a shared conceptualization of a given domain of interest [9]. Lai defines ontology as a means of enabling communication and knowledge sharing by capturing a shared understanding of terms that can be used both by humans and by programs [16]. It defined a complex relations map that requires to be formulated in an exhaustive and rigorous conceptual scheme to constitute a knowledge base by capturing a shared meaningful of terms. It leans on this perspective to show understanding information (semantic). According to [10], this form of knowledge is considered as an intermediate representation of a conceptualization that is more formal and structured than the natural language, but less formal than a formal language, which allows establishing a common language which can be understood and capture the accumulated knowledge. OWL is the widely accepted approach to standardize a language for ontologies. OWL ontology consists of Individuals, Properties and Classes. Individuals

(also known as instances) represent objects within a given domain. Properties state relationships between individuals or between individuals and data values. OWL class defines a group of individuals that share certain properties [8]. In this way, the data are referenced by metadata, under a standard normalized scheme which represent an abstract model of the real word. This mode of representation allows the interchange data following these standard schemes and, even, they can be modified and reused.

B. Ontology based multi-criteria decision making

In complex decision making methods the goal is to choose the best alternatives from a set of possible alternatives. The topic of decision making has been significantly spread in recent years. Generally, two types of approaches in the field of decision making can be distinguished: decision making based on quantitative and decision making based on qualitative models [17]. In quantitative models, the criteria values are numerical and continuous, whereas qualitative models consist of discrete criteria, whose values are presented by words rather than numbers [18]. Qualitative models provide a suitable representation, in many real life decision problems, such the information about criteria values are usually defined in the form of linguistic variables. Since the decision making internship assignment system presented in this paper is based on ontologies, it is based on qualitative multi criteria decision making. Our proposed ontology based decision making system enables universities a higher degree of internship assignment task automation since the problem solution can be obtained by reasoning upon the ontology. In our context, the ontology is introduced to guide the design of system, to supply it with semantic capabilities and to allow for defining an ontology guided internship assignment which provides intelligent matches between internship offers and candidates. Thus, in our ontology recommender system, the decision model is created from the existing ontology which describes the domain of student recruitment process in which the decision takes place. We refer to the ontology described in the next section as base ontology and the decision model.

II. RELATED WORK

In this section, we review the related work in e-recruitment and in ontology mapping by selecting the articles that are more relevant to our research. Many e-recruitment tools for employing candidates for a job have significantly spread in recent years. However, there has been relatively little research exploring semantic balanced matching systems between the concrete needs of the companies and employee's competencies. Thus, all the developed tools suffer from inadequate matching of candidates with job requirements [19]. Basic theory and mathematical tools are already available, but, the most complicated part in job matching process is the matching between the candidate's information and employers' requirement. To the best of our knowledge, no previous published work has applied a recommender system integrating semantic information related to the student competencies and interests, training courses and job

postings in order to draw a balance taking the requirements of companies across the skills of the students.

A lot of previous works have focused on extracting information about skill demands including the qualifications and skill requirements in the job market [12]. [13] uses text mining to extract and map skills listed in job postings to defined occupations. The work does neither implement an automatic pipeline nor an extensive evaluation to prove the proposed approach. [14] deals with skills required for Business Intelligence and Big Data jobs by finding the patterns to discover similarities and differences using the Latent Semantic Analysis (LSA) and Singular Value Decomposition (SVD). In [15], a cluster analysis in two phases was demonstrated to analyze job adverts based on current skill sets by applying cluster analysis with hierarchical agglomerative clustering and k-means. In [14], a skill system is established for generating occupational competencies with an automated approach using Named Entity Recognition (NER) and Named Entity Normalization (NEN) for raw texts.

Other several attempts were made efforts to automate recruitment process [20] by setting up a recommender system in order to match the right job with the right candidate. The recommender systems usually compare the collected data with similar data collected from others and calculate a list of recommended items for the user. To do so, recommender techniques such as content-based filtering, Rule-based filtering, Collaboration filtering and Hybrid filtering can be applied [2].

Related studies believe that interactions are important for recommendation [5] as they have a great impact on the candidate's job choice and employer's hiring decision. Some interaction-based recommendation systems, such as CASPER [7], make use of collaborative filtering to recommend jobs to users based on what similar users have previously liked. Hybrid systems are also applied to match people skills and jobs description offer by taking into account both the preferences of the recruiters and the interests of the candidates.

There are many various approaches are introduced with an objective to automate the recruitment process. [21] presents an SMS-based recommendation system for campus recruitment in China, which helps college placement office to match the companies and students with higher precision at lower cost. They are mainly focusing on profile matching and preference-list based two-sided matching for further recommendation. [22] proposed PROSPECT, a system for selecting candidates for recruitment. They exploit resumes to extract relevant aspects like competencies, experience in each skill, education details and past experience. [23] investigates and suggest the revised resume format that includes candidate personality assessment information for improving the effectiveness of candidate screening and selection.

There are many others e-recruitment systems that have been defined with an objective to speed-up and increase the efficiency of the recruitment process. In order to find the suitable candidates for job positions, these systems use different approaches like relevance feedback (Kessler et al., 2007; Yi et al., 2007), semantic matching (Mochol et

la., 2007), machine learning (Faliagka et al., 2012), natural language processing (Amdouni and Ben Abdesslem Karaa, 2010) and analytic hierarchy process (Faliagka et al., 2011) to automatically represent resumes in a standard format.

In order to improve the alternative education, we want to provide in this paper a semantic recommender system based on ontological models that allows us to better capture, analyze and use relevant semantic information for the exploitation and the simultaneous assignment of the internships to the students. Besides, our system will exploit the relationship between the demands of the missions and the student's profile who have followed a number of training courses and have some experience in finer detail. Consequently, the proposed system helps universities to satisfy both companies and students.

III. THE ONTOLOGY BASED MULTI CRITERIA RECOMMENDER SYSTEM

The proposed ontology based multi criteria recommender system enables to define decision models using ontology as the base construct. It structures decision models in such a way that the problem solution can be obtained by reasoning upon the ontology. Besides, the proposed recommender system is based on qualitative multi criteria decision making that is applied to the field of the ontology of the student assignment process. The ontological models defined in this paper capture more and more semantics from input models and provide us a support in our decision system for analyzing and assessing information by taking into account the evolution of the concrete needs of the company and the student's skills.

Our recommender system has three phases in screening candidates for internship recruitment. In the first phase, internship requirements are represented as ontology. In the second phase, the system captures all the information of the students related to their skills, personal information, work experience, current education and evaluation results and constructs their corresponding ontology models. In the third phase, we construct training programs ontology in which we collect all document for the features of the learning university programs like target competencies, modules, teaching units and the main goals of the learned module. Finally, we discuss the steps of the ontology based multi criteria decision making approach that enables to screen the eligible qualifications for an internship positions.

A. Phase1: internship requirements ontology

Nowadays, internship descriptions are written in form of free text using uncontrolled vocabulary. In contrast, semantic annotation of internship postings using concepts from a controlled vocabulary, based on Semantic Web technologies leads to have a standardize structure of mission's descriptions and consequently a better matching of student's skills and internship postings. The detailed mission is described in our previous work [24]. We present in this subsection the most relevant concepts in mission posting model as follows:

- *Competencies*: presents the competences required by the company. It is composed of two subclasses which are *Action* and *Domain-Action*.

- a) *Action*: presents the keywords (verbs) that describe the actions required to do at the internship.
- b) *Domain-Action*: presents the domain of the actions.

The compositions of the two instances of the action and the domain-action help us to specify the student's diploma required by the company.

- *Experience*: define the required student's experiences.
- *Activity-Area*: specify the area of activities required in the internship
- *Company*: define the name of company, the company importance and the number of employees.

B. PhaseII: Student profile ontology

The student profile is created by taking all information from the student. It consists of the classes' definitions based on shared attributes, emphasizing on information needs, access conditions, experience, competencies and knowledge. [24] presents the ULL student profile model in detail. In this subsection, we extract the most important information related to the student of ULL that are needed in our recommender system and then we construct our student profile ontology model.

- *Academic-information*: details the academic information such as the degree of the student and the actual academic year.
- *Administrative-information*: contains the personal details of the student such as its first and last name, number phone, address, email, nationality and its age.
- *Evaluation-record*: contains the information about the level-professional of the students i.e. the notes of its oral presentation, its quality of work and its behavior obtained during its internship in the company. Its evaluation-record is evaluated by the member of both company and university.
- *Candidate-record*: contains the information about the academic level of the student by evaluating him in some items such as quality of its experiences, its knowledge of the field of project management and monitoring process, overall rating of its curricula vitae, etc.
- *Interests*: they express the preferences of the students in terms of mission, location, salary and company.
- *Acquired competencies*: represents the competencies that are acquired in their previous learning courses.
- *Actual- competencies*: defined the skills in the process of being acquired in their actual diploma.

C. Phase III: The training course ontology

Training learning ontology model is important in the design of our recommender system because it allows us to make a link between students and companies and to define some rules and constraints. In fact, students upgrade their skills in their universities in order to reach the highest

level that would enable them to succeed in the professional world. We present in this subsection the most retrieval concepts of this ontology that are necessary for the understanding and the conceptualization of our system.

- *Teaching units*: define the objectives of the training and details the target competencies. They are represented by keywords.
- *Module*: represents a set of courses that must be taught to the students.
- *Course*: contains a finer level of competencies that are presented by keywords.

D. Phase IV: The steps of the ontology based multi criteria decision making method

As we described above, the internship assignment is a very complicate process. The main raison is that the companies want to have the best qualifications which correspond to their concrete needs. Generally, the top talents of students are selected by several companies. In contrast, some of students don't find the appropriate internship that interests them and some others don't find any internship and consequently they would not validate their degrees.

To resolve this problem, multi criteria recommended system must be deployed to assign an internship to every student. As we say above, this system consists of balancing of company's needs described in missions and student's capabilities

Our ontological models defined above allow us to perform semantic matching and provide a support for our complex problem-solving and facilitate knowledge modelling and reuse. There are therefore capable to analyze and assess our decision support system by taking into account the evolution of the concrete needs of the company and the student's skills.

1) Step1: Clustering the internship postings

To set up our recommender system, we first structure the missions' offers that are defined descriptively in a natural language into a formal annotation by instantiating the missions' concepts. This is can be done by extracting the most important keywords and introducing them into internship' concepts. These retrieval keywords are constructed and maintained by a domain expert. In our context, the concepts that we have taken into account to regroup the most similar mission are:

- Level of study requested
- Name of the company
- Activity-Area
- Age
- Descriptive of the requested tasks: Action and Domain-Action

We regroup the similar missions' instances into a same cluster in term of their structure and semantic annotation by applying the clustering algorithm *K-means* [26]. The distance function that is used in *K-means* to compute the distance between two mission m_1 and m_2 is calculated as follows:

$$\text{Distance-Mission}(m_1, m_2) = \sqrt{\sum_{n=1}^5 (w_n d_n^2)} \quad (1)$$

where d_n is the distance between different concepts of m_1 and m_2 and w_n is the weight associated to each distance. The definitions of all distances that calculate the similarity between two missions are given in table 1.

TABLE I. THE DEFINITION OF THE DISTANCES USED IN THE SIMILARITY FUNCTION OF K-MEAS

	<i>Noun</i>	<i>Function</i>	<i>Weight</i>
d_1	DistanceTasks(m_1, m_2)	ADW (descriptive1, descriptive2)	1
d_2	DistanceLevelStudy	Min-Max (level-study1, level-study2)	0,5
d_3	DistanceNameCompany	returns , if we have the same name of company 0, otherwise	1
d_4	DistanceActivityArea	returns , if we have the same Activity-Area of company 0, otherwise	1
d_5	DistanceAge	Min-Max (age1, age2)	0,5

$$\text{where Min-Max } (x_1, x_2) = \frac{\min(x_1, x_2)}{\max(x_1, x_2)}$$

To compute the similarity between the requested tasks which are defined by action and domain-action described in the internship posting, we use the cosine similarity of the unified approach for measuring semantic similarity ADW: *Align Disambiguate and Walk* [25]. ADW is an essential component of many Natural Language Processing applications. It operates at multiple levels all the way from comparing word senses to comparing text documents by using WordNet ontology based state-of-the-art semantic similarity that provides a rich network structure of semantic relatedness, connects senses directly with their hypernyms, and provides information on semantically similar senses by virtue of their nearby locality in the network. ADW leverages a common probabilistic representation over word senses in order to compare different types of linguistic data.

2) *Step2: The assigning of the new missions in the existing clusters*

We assign each new internship offer in the appropriate cluster by calculating the distance between the new mission and the centroids of each cluster. We use the similarity function (1) to compute the distance between the two mission m_x and c_i where m_x is the internship offer to be affected in the correspondent cluster and c_i is the mission that represents the centroid of the cluster.

3) *Step3: The selection of the nearest missions for the new internship offer in the cluster*

For each new mission assigned to the appropriate cluster, we select 20% of the nearest internship postings by using the similarity function (1) in order to determine the students who have succeeded these missions and consequently it help us to know the type of profiles that correspond to the new mission. The nearest missions are called the ClosestOldMission of the new mission.

4) *Step4: The selection of the old students satisfying the ClosestOldMissions*

We select in this step the individuals of the ClosestOldMissions satisfying the necessary and sufficient conditions for success. Supposing the criterion subclasses Satisfactory of the class ClosestOldMissions

are: ExcellentStudent, Very-GoodStudent, appropriateStudent and PassableStudent. All the individuals belonging to the subclass ClosestOldMission are retrieved. We called this group of students the ClosestOldStudents. This group allows us to predict the type of profiles that correspond to a new mission.

5) *Step5: The determination of the decision criteria for the selecting of the students candidates.*

We identify in this step the criteria that influence the student selection decision. In order to make a decision, evaluation of candidates in the classes of alternatives according to the criteria is required. Therefore, the condition that the criteria elements have to satisfy is that they directly or indirectly describe the individuals in the class of alternatives.

We define the following criteria:

a) *CR1: The similarity between the candidate's profile and the ClosestOldStudents' profiles*

We compute the similarity between the candidate' profile and the students' profiles of the ClosestOldStudents of the new mission M_x with the following:

$$\text{Sim-Students } (S, N) = \frac{\sum_{i=1}^N \text{SimProfiles } (S, S_i)}{N} \quad (2)$$

Where N is the size of the ClosestOldStudents and the similarity SimProfiles between two students S_1 and S_2 is defined as follows:

$$\text{SimProfiles } (S_1, S_2) = \sqrt{\sum_{n=1}^6 P_n W_n^2} \quad (3)$$

Where P_n are the different parameters that we have taken into account to compute the similarity between the profiles of the students and W_n is the weight that reflects the importance of each parameter.

All the defined parameters are presented in the bellow table.

TABLE II. THE DEFINITION OF THE PARAMETERS USED TO COMPUTE THE SIMILARITY BETWEEN STUDENTS' PROFILES

	<i>Noun</i>	<i>Function</i>	<i>Weight</i>
P ₁	Notes-Evaluation	Cosine similarity (NotesEv1, NotesEv2)	7
P ₂	Training taken by students	Returns: 1, if T _C and T _P are the same 2/3, if T _C are the same and T _P are different 1/3, if T _C are different and T _P are the same 0, if T _C and T _P are different	5
P ₃	Notes-Training	Cosine similarity (NotesTr1, NotesTr2)	5
P ₄	Interests	Cosine similarity (Interests1, Interests2)	3
P ₅	Age	Min-Max (age1, age2)	2
P ₆	Localization	Min-Max (kilometerNb (local1, localM _x), kilometerNb (local2, localM _x))	1

- a. T_C: training of the current year
b. T_P: Training of the previous year

b) *CR2: The similarity between the requested competences in the new mission and the formation and competences of the candidate*

In the multi criteria recommender system, we define the similarity between the competencies required in the internship offer M_x and the formation F taken by the students by computing the semantic similarity between them and taking into account the notes of the students in each formation. Therefore, the similarity function is defined as below:

$$SimFormation(F, M_x) \quad (4)$$

$$= \sum_{i=1}^N SimUE(T, M_x)_i Avg_i (ADW(items_x, items_{FOi}) + ADW(items_x, items_{FCi}) + ADW(items_x, items_{FTi}))$$

where:

N is the number of the formations

Avg_x is the average obtained by students in X

$Items_y$ is the extracted keywords from the descriptive of Y ;

$Y = \{X: \text{mission, FO: Formation objective, FC: Pedagogical training, FT: expected work in the missions}\}$

$SimUE(T, M_x)$ is the similarity function between teaching units T of a formation and the mission M_x , is represented in (5).

$$SimUE(T, M_x) \quad (5)$$

$$= \sum_{j=1}^k SimM(M, M_x)_{Tj} Avg_j ADW(items_x, items_j)$$

where:

k is the number of the teaching units and $SimM(M, M_x)$

is the similarity function between the module M and the mission M_x which is denoted as follows:

$$SimM(M, M_x) = \sum_{i=1}^L Avg_i ADW(items_x, items_j) \quad (6)$$

c) *CR3: The adequacy between the company' constraints and the candidate' profile*

The sub criteria considered are given in table 3.

Each sub criteria is associated with a weight that refers to its importance.

TABLE III. THE DIFFERENT SUB CRITERIA TAKED INTO ACCOUNT TO COMPUTE THE ADEQUANCY BETWEEN MISSIONS' CONSTRAINTS AND STUDENT' PROFILE

	Noun	Function	Weight
S_1	The required level of study	Returns: 1, if the required level of study of the student $\geq L$ 0, otherwise	7
S_2	Age	Mix-Max (age-student, age-mission)	5

L. The required level of study defined by the company

The adequacy between mission' constraints M and student' profile S is computed in (7).

$$Adeq = \frac{\sum_{i=1}^2 S_i W_i}{2} \quad (7)$$

d) *CR4: The adequacy with the interests of the candidate*

The similarity function is calculated with the following:

$$SimInterests = \sqrt{\sum_{i=1}^4 (I_i W_i)^2} \quad (8)$$

where I is the interests and W is the weight affected to each interest. Table 4 defines the different interests which are considered in the proposed multi criteria making recommended system.

TABLE IV. THE DIFFERENT SUB CRITERIA CONSIDERED TO COMPUTE THE ADEQUANCY OF CANDIDATES' INTERESTS WITH THE INTERNSHIP OFFERS

	Noun	Function	Weight
--	------	----------	--------

	Noun	Function	Weight
P_1	Activity-Area	nbAAC/maxnbAA	7
P_2	Occupation	nbOC/maxnbO	5
P_3	Company size	Min-max (Size-Company, wished-size)	5
P_4	Salary	Min-max (Salary-company, wished-salary)	3

- nbAAC: The number of common activity area of the proposed activity area by company and the activities that interest the candidate
- nbAA: the number of activity area
- nbTC: the number of common occupations of the proposed occupations in the internship and the occupations that interest the candidate
- nbO: the number of occupations

e) *CR5: The localization*

The localization is defined as below:

$$Loc(X, S) = d \quad (9)$$

where

$d = 1$, if $L \leq 4$ km

$d = 2$, if $4 < L \leq 10$ km

$d = 3$, if $10 < L \leq 20$ km

$d = 4$, if $20 < L \leq 50$ km

$d = 5$, if $L > 50$ km

and $L = \text{nbkilm (GS, GX)}$; GS and GX represent GPS coordinates of the localization of the student S and the Company X respectively.

6) *Step6: The application of the AHP for the selecting and ranking the student in a given mission*

The procedure that determines how best to evaluate alternatives or to decide which alternative is preferred to another is known as decision rule. It integrates the data on a set of alternatives and decision-maker's preferences into an overall assessment of each alternative. The process of applying the decision rule is concerned with the appropriate combination of the relevant criteria to determine the overall evaluation scores (ratings and rankings) for the decision alternatives.

The AHP method is a decision rule. It has been identified as an important approach to multi-criteria decision-making problems of choice and prioritization. Its extensive application is due to its simplicity, ease of use, and flexibility.

The AHP decision problem is structured hierarchically at different levels, each level consisting of a finite number of decision elements. The top level of the hierarchy represents the goal; one or more intermediate levels embody the decision criteria and sub-criteria while the lowest level is composed of all potential alternatives. The relative importance of the decision elements (weights of criteria and scores of alternatives) is assessed indirectly from comparison judgments' during the second step of the decision process. The domain expert is required to provide his/her preferences by comparing all criteria, sub-criteria and alternatives with respect to upper level decision elements. The values of the weights and scores are elicited from these comparisons and represented in a decision table. The last step of the AHP aggregates all local priorities from the decision table by a simple weighted sum. The global priorities thus obtained are used for final ranking of the alternatives and selection of the best ones that are used as crisp values for final ranking of alternatives.

IV. IMPLEMENTATION OVERVIEW

We have chosen Protégé tool based on the ontological language in order to implement our ontological models, due to its extensibility, quick prototyping and application development. Protégé ontologies are easily exported into different formats including RDF Schema, Web Ontology Language (OWL) which is defined as a language of markup to publish and to share data using ontologies in the WWW and Top Braid Composer. Therefore, we have used the reasoning Pellet to check the consistency of the SWRL rules and to execute them.

To establish the ontology based multi criteria making recommended system, we have developed a java implementation in order to exploit the knowledge base created by instantiating our ontological models. This program allows the user to query the news and view the knowledge base. It uses the framework Jena which provides integrated implementations of the W3C Semantic Web Recommendations, centered on the RDF graph for manipulating and reasoning with the OWL ontologies. For querying, it employs SPARQL and tSPARQL [11], which adds time functionalities to the queries. Then, we will use the Weka framework to implement our clustering and optimization algorithms. Finally, the results will be returned in a web interface.

V. CONCLUSION

In this paper, we proposed an ontology based multi criteria recommended system to guide the assignment internship process. We described the ontological models, the student profile, training course and internship posting, used within our approach which provides us means for semantic annotation. Using controlled vocabularies, in contrast to free text descriptions, results in a better machine process ability, data interoperability and integration. Moreover, having internship offer and student profiles semantically annotated, enables us to perform semantic matching which significantly improves query results and delivers a ranked list of best matching candidates for a given internship position.

ACKNOWLEDGMENT

This publication was made possible by NPRP grand # NPRP 7-1883-5-289 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

VI. REFERENCES

- [1] A. Sieg, B. Mobasher, R. Burke, "Ontological user profiles for personalized web search", 5th Workshop on Intelligent Techniques. AAAI, 2007.
- [2] Q. Gao, J. Yan, M. Liu, "A Semantic Approach to Recommendation System Based on User Ontology and Spreading Activation Model", Network and Parallel Computing International Conference, 2008.
- [3] N. Bolaian et al, "A Model for a Collaborative Recommender System for Multimedia Learning Material", Lecture Notes In Computer Science, Vol. 3198, pp. 281-288, 2004.
- [4] L. Ramos, R. Gil, «Propuesta de Sistema de Información para Apoyar la Gestión de la Educación a Distancia », Conferencia Iberoamericana en Sistemas, Cibernética e Informática, 2007.
- [5] B. Mobasher , X. Jin, Y. Zhou, "Semantically Enhanced Collaborative Filtering on the Web", Lecture Notes in Computer Science, 2004.
- [6] J. Pan, B. Zhang, S. Wang, G Wu, D. Wei, "Ontology Based User Profiling in Personalized Information Service Agent, Computer and Information Technology", 2007, 7th IEEE International Conference, 2007.
- [7] Rafter, R., Bradley, K., and Smyth, B. 2004. "Automated collaborative filtering applications for online recruitment services". In Proceedings of the International Conference on Adaptive Hypermedia and Adaptive Web- based Systems. 1892, Trento, Italy, August. 2000.
- [8] W3C, OWL Web Ontology Language Overview, 2004. <<http://www.w3.org/TR/owl-features>>.
- [9] Y. Sure, R. Studer, " Semantic web technologies for digital libraries", in: Library Management, vol. 26, nos. 4 and 5, Emerald Group Publishing Limited, UK, 2005, pp. 190-195 (Special Issue: Semantic Web).
- [10] M. Uschold, M. King, "Towards a Methodology for Building Ontologies", Workshop on Basic Ontological Issues in Knowledge Sharing, 1995
- [11] F. Frasincar, J. Borsje, and L. Levering. "A Semantic Web-Based Approach for Building Personalized News Services", International Journal of E-Business Research, pp. 35-53, 2009.
- [12] Zhao, M., Javed, F., Jacob, F., McNair, M.: SKILL, "A system for skill identification and normalization", In: Proceedings of the twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, Texas, USA. pp. 4012-4018 , January 25-30, 2015.
- [13] Wowczko, I.A, "Skills and vacancy analysis with data mining techniques", Informatics, vol. 2, pp.31, 2015.
- [14] Debortoli, S., Muller, O., vom Brocke, J., " Comparing business intelligence and big data skills", Business & Information Systems Engineering, vol. 6, pp. 289-300,2014.
- [15] Litecky, C., Aken, A., Ahmad, A., Nelson, H.J., "Mining for computing jobs", IEEE Software, vol. 27, pp. 78-85,2010.
- [16] L.F. Lai, "A knowledge engineering approach to knowledge management", Information Sciences, vol. 177, pp. 4072 - 4094, 2007
- [17] L. Martí nez, J. Liu, D. Ruan, J.-B. Yang, "Dealing with heterogeneous information in engineering evaluation processes", Information Sciences, pp. 1533-1542, 2007
- [18] M. Bohanec, B. Zupan, V. Rajkovic, "Applications of qualitative multi-attribute decision models in health care", International Journal of Medical Informatics, pp. 191-205, 2000
- [19] C., Heese, R., Mochol, M., Oldakowski, R., Tolkendorf, R. and Eckstein, R., "The impact of semantic web technologies on job recruitment processes", Proceedings of International Conference on Wirtschaftsinformatik, Banberg, Germany, pp.137-138, 2005
- [20] Lee, I., "An architecture for a next-generation holistic e-recruiting system", *Communication ACM*, Vol. 50, No. 7, pp.81-85, 2007
- [21] X., Wu, L., Li, C. and Huang, M., "SMS-based mobile recommendation system for campus recruitment in China", 10th International Conference on Mobile Business, pp.152-157, 2011
- [22] Singh, A., Rose, C., Visweswariah, K., Chenthamarakshan, V. and Kambhatla, N., "PROSPECT: a system for screening candidates for recruitment", CIKM'10 Proceedings of the 19th ACM International Conference on Information and Knowledge Management, pp.659-668, 2010.
- [23] Wright, E and Domogalski, T., "Improving employee selection with a revised resume format", Business Communication Quarterly, Vol. 74, No. 3, pp.272-286, 2011.
- [24] A. M'baya, J.Laval, N. Moalla, Y.Ouzrout and A. Bouras, "Ontology based system to guide internship assignment process", in the 12 International Conference on Signal image technology and internet based systems in Naples, 2015
- [25] Pilehvar, M. T., Jurgens, D., & Navigli, R., "Align, disambiguate and walk". In Association for Computational Linguistics (ACL), 2013
- [26] Wagstaff, K., Cardie, C., Rogers, S., & Schrödl, S., "Constrained k-means clustering with background knowledge", In ICML, Vol. 1, pp. 577-584, 2001.