Web Content Mining Framework for Discovering University Formations' Compatibility with the Market Needs

Camelia Elena CIOLAC
The Bucharest Academy of Economic Studies, Romania
E-mail: ciolac_c@yahoo.co.uk
Florica LUBAN
The Bucharest Academy of Economic Studies, Romania
E-mail: florica.luban@man.ase.ro
Răzvan Cătălin DOBREA
The Bucharest Academy of Economic Studies, Romania
E-mail: razvan.dobrea@man.ase.ro

Abstract

This paper aims to provide a decision tool for both academia and business in order to better reach convergence between the qualifications offered and demanded for fresh graduates. The study discovers the IT market's request of capabilities by applying web-mining techniques on the main Romanian job web portals. Then, the study gathers information about the specific qualifications provided to students in both undergraduate and master computer science programs in five representative Romanian universities. Based on these findings, we propose a model to evaluate the degree of alignment of these two sets. The model is useful for all stakeholders: universities, companies including small and medium size enterprises, as a tool of career planning and it can be adapted to assess the best mobility program for students, as further discussed in the paper.

Keywords: *knowledge discovery, web content mining, e-recruitment, software agents, IT education*

JEL classification: D83, L86, I23, J24

Introduction

The paper aims to propose a automated web content mining framework to evaluate the compatibility between university formations and market qualification needs. The proposed approach brings many advantages that contain, but are not restricted to the following:

- a comprehensive mechanism of acquiring knowledge from the recruiting websites;
- a reliable and objective instrument to facilitate education alignment to market qualification needs;
- a guideline for companies that organize internships to better understand the educational background of student candidates;
- a framework to asses the opportunity of students and teachers mobility between universities;
- a tool for career planning for both enrolled and prospective students.

Based on this framework, the paper then presents a study in the IT (Information Technology) field carried out in Romania in July-August 2009.

A considerable number of research papers addressed the topic of business needs – university curricula convergence. Reeves and Bussom (1980) conducts a survey among managers of information systems activities in Fortune 500 organizations regarding the „educational output objectives" they desire from graduates of business universities. Forgionne (1992) underlines the fact that each university program provides „partial, but not complete and integrated, information science education and research" and analyzes the education and research gaps to practice. Richards & Pelley (1994) makes „an attempt to identify curriculum components which can be considered valuable" by carrying out a survey among employed graduates in IS. Lee et al. (2002) measure the gap between the IT skills considered important by academics versus practitioners. Wenger (2003) analyzes the trend in programming languages and puts into light the industry's recommendations for changing the computer science curricula in universities. Kitchenham et al. (2005) assess „extent to which the education delivered by four UK universities matches the requirements of the software industry" using the surveys method. Li et al. (2008) evaluate the gap between university formations and e-business needs of qualifications through a comparative research between the US and Taiwan. Cyrenne & Grant (2009) discuss the factors that influence the reputation or prestige of universities, including the signals of quality that the universities deliver. Ono (2008) examines the relationship between university prestige and socioeconomic achievement in Japan and underlines the difference between national universities' graduates and private universities' graduates.

Regarding the criteria that can be used to evaluate a graduate program, the literature provides several approaches. Burgoyne (1975) proposes the following main aspects: „(a) Access to and progress in successful careers; (b) Flexibility and mobility in careers; (c) Ability to take overall view of problems; (d) Personal confidence; (e) Awareness of own aspirations and career fit; (f) Frustration from unused skills and unfulfilled expectations; (g) High salaries; (h) General reasoning and problem solving ability; (i) Economic/commercial understanding; (j) Understanding of organization and human behavior; (k) Useful work, career and personal contacts; (l) Social skills". Alexander (2000) states the importance of curriculum standards in order to obtain performance in computer science education. Luban (2005) underlines the fact that academia must continuously adapt their curricula to the society's needs as well as to increase the quality of knowledge that future graduates possess. Still, the author admits that technical knowledge is only partially obtained in university and on-the-job training is needed to improve the employee's compliance to the job technical requirements. In (Nicolescu, 2009), it is shown that a new category of changes based on knowledge is developing and their efficacy depends on the quality and degree of capitalization of knowledge generated inside and outside the organisation. Johnes (2006) analyzes teaching efficiency by using data envelopment analysis (DEA) and considering Economic graduates from UK Universities.

In terms of training on the job, Salas-Velasco (2009) elaborates a regression model that estimates the probability that a fresh graduate receives training on the job, taking into consideration the occupation, the economic sector, the employer company's size.

One of the main concerns for the Romanian National Council of Adults' Professional Training is the definition of the core competencies that best describe the modern economy's jobs. For the IT sector, CNFPA, identified the following occupations that are of interest for most of the industries, for both large organizations, such as banks, and for SMEs (small and medium enterprises): Database administrator, Application administrator, Network administrator, Analyst, IT consultant, Web designer, IT manager, DTP designer, Software engineer, Computer operator, Programmer assistant, IS programmer, IS designer, computer-aided design specialist, Security specialist (CNFPA, nd).

According to Ratiu-Suciu & Stoica (2004), the process of knowledge market modeling involves different stages, which could lead to a market balance. Let Qof be the knowledge offered on the market and Qc the knowledge requested on the same market. The equations (1) and (2) calculate the disequilibria convenient for the offer gd^{of}, respectively favorable for the demand gd^{c}.

$$gd^{of} = (Qof - Qc) / Qof \tag{1}$$
$$gd^{c} = (Qc - Qof) / Qc \tag{2}$$

The knowledge market equilibrium is characterized by the relation (3).

$$gd^{of} = gd^{c} = 0 \tag{3}$$

In (Ratiu-Suciu & Stoica, 2004), it is shown that the decisions related to each disequilibria situation are:

- should the knowledge offer exceed the demand, then companies increase their Research & Development budget in order to take advantage of the under-utilized knowledge available on the market;
- should the market register a knowledge deficit, companies reduce investments in knowledge transfer and increase investments in technology transfer.

Reported to these studies, the present paper proposes an innovative framework of automatically measuring the gap between industry and academia in terms of IT competencies through web content mining.

1. Preliminary considerations

In order to automize the knowledge acquisition through web content mining, we use software information agents as described in section 3 of the paper.

The target websites that are subject to web content mining are split in two main categories:

- University websites;
- e-Recruitment websites.

Each of the two categories has a set of features that must be considered especially in the process of software agents' training.

University websites features

In order to acquire the education plans from a University website, two scenarios must be considered:

- the education plans are gathered and directly accessible at the root node of the University website file tree;
- each faculty / department of the university provides its educational offer on its own website, whose link can be accessed from the main University website. Often, faculty / department URL cannot be guessed given the university's URL, because even if hosted on the same web server, can have different formats.

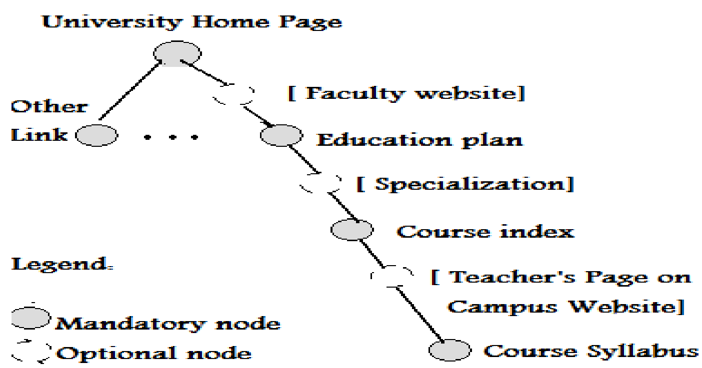

Figure 1 Website structure scenarios for university education plan

A rarely encountered situation is the one where course syllabus is provided in teachers' personal page and not via the education plan. In Figure 1 are represented the main scenarios that software agents should be capable of dealing with appropriately.

e-Recruitment websites features

These websites provide generally two means of achieving the index of jobs from an industry: through a search form or through link cascade navigation.

Considering the fact that e-Recruitment websites organize their categories upon different criteria (e.g. location->industry->jobs, industry->location->jobs, industry->subindustries->jobs), training a software agent to adapt to this variety of scenarios is more complicated than training it to get the desired results through form input and submission.

A typical search input form provided by an e-Recruitment website consists of the following elements:

- a job name – as a text field
- an industry – as a dropdown list
- a career level – as a dropdown list (optional)
- a city name – as a dropdown list
- a submit button.

Optionally it may contain choice for country as a dropdown list or for time employment type.

2. Framework description and implementation details

The proposed framework facilitates knowledge acquisition and processing following the steps depicted in Figure 2 and described in the following paragraphs.

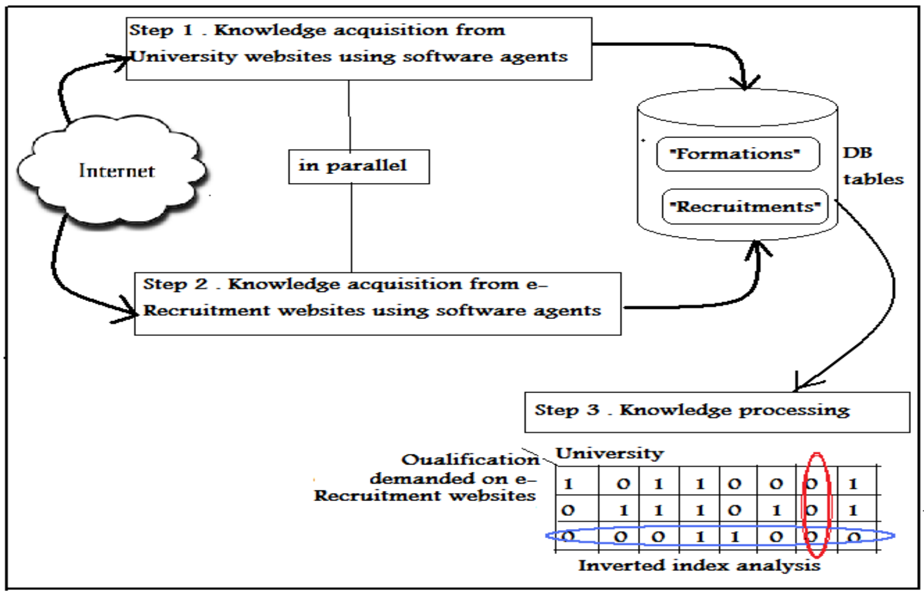

Figure 2 Workflow in the web content mining framework

Step 1. Knowledge acquisition from University websites using software agents.

Given the limited number of universities in a country and the static structure of their websites, it is little probable that the website would be restructured significantly in the future. Therefore a database of agent behavior for each university website proves to be an appropriate approach. Thus, a record in the table for agent behavior in this case should have the structure described in Figure 3.

UId	*LastUpdate*	*RootURL*	*BaseURL*				

Figure 3 The nested table of parameters

For example, *BaseURL='http://www.csie.ase.ro/Licenta/ Default.aspx?view='* and the parameter list contained in the nested table has the varchar2 elements: *'6'*, *'7'*, *'8'*. Generally, the nested table should contain the variable part of the URL, that appended to the *BaseURL* generates a valid address of an education plan.

RootURL represents the root from which the worker begins its search activity. The software implementation is based on the Replicated Workers model. Periodically, the tasks pool is populated with tasks of University webpages mining. Each task *ti* is associated with an unique *UId* (University ID), *uidi*.

A worker *wi* that explores a University *uidi* website firstly checks if the information from the „Formations" table of the database is synchronized with the actual contents of the University *uidi* website's education plans. In order to accomplish this step, a HTTP Head message is submitted and, at response arrival, the worker compares the value of LastUpdate corresponding to its *UId* with the value of Last-Modified from the HTTP Response. In case the two values are equal, the task is considered solved and the worker thread returns to pick another task from the task pool.

Contrarily, if the two dates differ, the worker *wi* iterates through all the URLs that can be formed by appending a nested table item to the end of the *BaseURL*. For each Analyzed URL *urlj* built in this manner, the worker *wi* submits a HTTP Get request and receives in return an index of courses *Ic* available for that specialization / education plan.

The worker *wi* then activates a software agent *ai* to mine the syllabus of each course. The agent receives as an input the courses index *Ic*. For each entry *ci* in the courses index *Ic*, the agent *ai* tests whether it has associated an < a href=..> Label . If the entry *ci*'s structure in the courses index *Ic* presents more than one link *lj, j=1,…,n* anchors, the agent will process them iteratively, excepting the „mailto" elements.

For each link *li* processed ,the agent opens the referred web document and searches for key words, such as „contents", "description", „objectives", "references".

The result of the data mining process for a course *ci* syllabus is stored in the „Formations" table of the database, with respect to maintaining the table without duplicate entries.

Challenges:

(A) An exceptional situation arises when this algorithm is run for the first time that is when the nested table of parameters for each University ID is empty. In this case, before activating the mining agent , the worker should get the set of all links available from the *RootURL* and decide which one of them references education plans (that is, contains keywords like „education plan", "course name" and <table> or / HTML tags).

(B) Another particular situation can arise in the course *ci* syllabus mining process when the web server answers to the agent's HTTP Get request with an error code of 301 (Moved Permanently), 404 (Not Found), 500 (Internal Server Error) etc. In this case, the agent should be temporary suspended and the worker should carry out the procedure described above in paragraph (A).

In conclusion, at the end of this step, the system acquired the necessary knowledge about the educational offer of all accredited universities in a country.

Step 2. Knowledge acquisition from e-Recruitment websites using software agents.

In this case, the strategy is slightly different. Each agent *ai* is assigned an industry name *In* and a set of e-Recruitment websites *WSi*={urlj | urlj is URL of an e-Recruitment site}. For each *urlj* in *WSi*, the agent *ai* retrieves the associated web document. Parsing the retrieved web document, the agent *ai* localizes the search input area and its composing items.

Suppose that the agent *ai* identifies the HTML items that form the search area at lines *l1, ..., ln* in the web document. It will store in a local versatile memory a vector of tuples:

(*li*, *input_label, input_type, options*),

where *li* represents the line number, *input_type* stands for the type of input element and *options* stores the vector of options in a dropdown list item.

The agent *ai* fills the text field with the industry name *In* and, if an industry dropdown list provided, selects the option that resembles to the searched industry (like '%'||*In*||'%') .

The index of available jobs for the entered selection is processed iteratively. For each job *j* in the result set, the agent verifies if the date of job *j*'s publishing belongs to the current university year. This supplementary test is necessary in order to map the timeline of market demand for qualifications to the university calendar of studies (e.g. from October 2008 until July 2009) for a pertinent analysis.

The job *j*'s required capabilities (qualifications) and job *j*'s location are stored in the table „Recruitments" of the database.

In case the job *j*'s required qualifications are too general and no specific capability is mentioned , the job *j*'s processing can return without any DML operation onto the database.

In case an advertised job has multiple employment locations, in different company branches, a record for each location should be inserted in the „Recruitments" table of the database.

An aggregate view „AggrRecruit" of the „Recruitments" table of the database is build, where for each qualification needed, a frequency of demand is computed, no matter the jobs' location. By ordering the „AggrRecruit" view's records descendant on the frequency field, we obtain valuable knowledge of the most demanded qualifications in the *In* industry in our country.

At the end of this step, the system acquired the necessary knowledge about the qualifications needed in practice, on the market.

Step 3. Knowledge processing.

Given the industry *In* that we analyze, let L={Loc | Loc= city in a country}

The previous *Step 1* provided the sets:

US = {*UId* | *UId*=ID of acredited University}

UL = {(*UId*, *ULoc*) | *UId* in US, *ULoc* in L University city location}

USC = {(*UId*, *SName*, *Q*) | *UId* in US, *SName* = specialization name, *Q* = qualification}

Furthermore, the previous *Step 2* provided the sets:

J = {(*Job, JLoc, Q*) | *JLoc* in L job location, *Q*= qualification}

JF = {(*Q*, *Fr*) | *Q* = qualification, *Fr*=frequency of demand for the qualification Q}

Let pr_{xi} (x1, x2, …, xn) = xi, be the regular projection function.

Using the projection function pr we can obtain the total set of qualifications (QS) offered or demanded in the industry *In* defined by (4).

$$QS = (U_{(x \in USC)}\ pr_Q\ x\ )\ U\ (U_{(y \in JF\ )}\ pr_Q\ y\ ) \qquad (4)$$

The analysis discovers the following knowledge:

K1) At a country level, the universities curricula (education plans) alignment to market qualification needs.

Let OQu be the set of qualifications offered by the University *u*, having the form (5).

$$OQu = U_{(x \in USC \text{ and } pr_{UId}\ x = u)}\ pr_Q\ x \qquad (5)$$

We build a binary function (6).

$$f: US\ x\ QS \text{ -> } \{0,1\} \qquad (6)$$

where $f(u, q) = \begin{cases} 1 \text{ iff } OQu \cap \{ q \} <> \emptyset \\ 0 \text{ else} \end{cases}$

Based on the binary function f, we build the inverted index as a binary matrix M defined by (7).

$$M= \begin{array}{c|c|c|c} \backslash\ UId & u1 & ..\ uj\ .. & um \\ Q\ \backslash & & & \\ \hline q1 & & & \\ \hline ..\ qi\ .. & & M[qi][uj] & \\ \hline qn & & & \end{array} \qquad (7)$$

where M[qi][uj] = f(uj, qi).

Suppose we want to obtain the university ID for all universities that prepare fully prepared candidates for a job j1 in J set. The following steps should be performed:

Step s1. Obtain the set of qualifications needed for job j1, denoted as QNj1.

$$QNj1 = U_{(x \in J \text{ and } pr_{Job}\ x = j1)}\ pr_Q\ x \qquad (8)$$

Step s2. For all $q \in QNj1$, we apply a logical AND operation on the corresponding lines of the M matrix. The result is used in direct access to discover those UId columns that satisfy the condition of preparing all the qualifications $q \in QNj1$.

Step s3. A deeper analysis could then consider inside each university *u* from the result set, whether there is possible, by following a certain specialization, to attend courses that provide the qualifications QNj1. This second analysis is important because there is a possibility that two different specializations of university *u* prepare the qualifications QNj1' and QNj1'' so that their reunion at the university level results in the set QNj1. Still an enrolled student cannot be fully prepared for job j1 in university *u* unless he/she follows both specializations.

Another important aspect that can result from this analysis is the extremes acknowledgement. Considering matrix M's columns, we can easily observe if there exists universities with old-fashioned education plans that provide qualifications that are not demanded on the market. Such an atypical university would have 0 elements on more than 80-90 % of its lines.

In an analogue manner, we can observe the atypical qualifications that are not widely provided in the higher education system. Generally, these atypical qualifications can be divided in two categories:

- new qualifications q, demanded by emergent technologies or by the knowledge-based economy, $q \in QS \setminus (U_{(x \square USC)}\ pr_Q\ x)$;
- old-fashioned qualifications q, that are no longer demanded on market, but that continue to be taught in a restricted set of universities, $q \in QS \setminus (U_{(y \square JF)} pr_Q y)$.

K2) At a local level, the qualifications available in a geographically delimited area.

In addition to the two dimensions considered in matrix M, by taking into account the spatial dimension of data , we obtain a 3-dimensional model.

K3) Similitude between universities' educational plans can be put in evidence.

Let us consider the distance between two universities as (9):

$$d(ui, uj) = < M[.][ui]\ \mathit{XOR}\ \ M[.][uj] >_0 \qquad (9)$$

meaning the number of identical elements on corresponding positions in columns ui and uj in binary matrix M . The notation $< x >_0$ means the number of bits of value 0 in the bit array x.

This similitude can be the basis for optimizing knowledge transfer in mobility programs for teachers and students, if the analysis is carried out at an international level.

K4) Optimization of investments in internships and employees recruitment**.**

Companies can find a valuable instrument in the process of HR decision making. Risks of hiring unsuitable fresh graduates diminish considerably when the company can evaluate by use of this framework whether an university fresh graduate is likely to have the necessary background for developing a career in that industry.

3. A Case Study in the Romanian IT Sector

In order to prove the applicability of the framework described in section 3, let us consider the following analysis in the Romanian IT sector.

The study was carried out in Bucharest, the capital of Romania, and consisted of analyzing the compatibility between university formations provided by the five representative universities located in Bucharest and the demand for qualifications expressed through e-Recruitment sites by IT companies activating in Bucharest.

The universities whose Bologna educational plans underwent the analyzing process are:

- An university with economic background Uni1
- A technical university Uni2
- A multidisciplinary public university Uni3
- An private university with international affiliation Uni4
- A multidisciplinary private university Uni5

Among these universities, Uni1, Uni2 and Uni3 are public education institutions, while Uni4 and Uni5 are private accredited universities.

In the following paragraphs, the following notations are used: Xu – the undergraduate studies in IT in university X, Xm – the master studies in IT in university X (where $X \in \{Uni1,Uni2,Uni3, Uni4,Uni5\}$).

The first analysis is a granular approach to evaluate the degree in which each university level of education (undergraduate/master) satisfies the qualifications needed on the IT market in their geographic area.

The qualifications needed on the actual IT market were collected from the main e-Recruitment websites in Romania (specified in bibliography). Queries had constraints in terms of industry and location (according to the theoretical framework described in section III of the paper).

In the matrixes presented in the following paragraphs, we use the aggregate columns:

T_u=Uni1u+Uni2u+Uni3u+Uni4u+Uni5u

T_m=Uni1m+Uni2m+Uni3m+Uni5m

T_edu=T_u+T_m

T_need = sum of distinct jobs that request the qualification, as resulted from the e-Recruitment websites.

We built the binary matrix M and obtain the following results:

R1) A first cluster of qualifications that are widely requested on the actual IT market and that are also contained in the major part of the universities formations (Figure 4).

In this cluster fall the qualifications that generally form the core of IT formations and represent the degree in which universities managed to adapt their education curricula to the real IT market needs. In this case, the number of ones on a formation column gives the degree to which the specific formation is adapted to the IT market demand dynamics.

Qualification	Uni1u	Uni2u	Uni3u	Uni4u	Uni5u	T_u	Uni1m	Uni2m	Uni3m	Uni5m	T_m	T_edu	T_need
OPERATING SYSTEMS	1	1	1	1	1	5	0	0	0	0	0	5	16
ORACLE	1	1	1	1	1	5	0	0	0	0	0	5	11
JAVA	0	1	1	1	1	4	1	0	1	0	2	6	8
COMPUTER NETWORKS	0	1	1	1	1	4	1	0	0	0	1	5	8
MYSQL	0	0	0	1	1	2	0	0	0	0	0	2	7
JAVASCRIPT	0	0	1	0	1	2	0	0	0	0	0	2	7
C++	1	1	1	1	1	5	0	0	1	0	1	6	6
XML	1	1	1	0	1	4	0	0	1	0	1	5	6
PHP	0	1	1	1	1	4	0	0	0	0	0	4	6
OOP	1	1	1	0	1	4	0	0	0	0	0	4	6
TCP/IP	0	1	0	0	0	1	1	0	0	0	1	2	6
SECURITY	0	0	1	0	1	2	1	1	0	0	2	4	4
PL/SQL	0	1	1	0	0	2	1	0	0	1	2	4	4
ALGORITHMS AND DATA STRUCTURES	1	1	1	1	1	5	0	0	1	0	1	6	3
UML	0	1	1	0	1	3	1	0	1	1	3	6	3
ERP	1	1	0	1	0	3	1	0	0	0	1	4	3
DBMS	0	0	1	1	1	3	0	0	0	0	0	3	3
DATABASE DESIGN	1	1	1	0	0	3	0	0	0	0	0	3	3
VISUAL BASIC	0	0	1	1	0	2	0	0	0	0	0	2	3
ASP.NET	0	0	1	0	0	1	1	0	0	0	1	2	3
STRUTS	0	0	0	0	0	0	1	0	1	0	2	2	3
C#	1	0	1	0	1	3	0	0	0	0	0	3	2
DATABASE MANAGEMENT	0	0	1	0	0	1	1	1	1	0	3	4	2
JSP	0	0	1	0	0	1	1	1	1	0	3	4	2
ASSEMBLER	0	1	0	0	0	1	1	0	0	0	1	2	2
PROJECT MANAGEMENT	0	0	0	0	0	0	1	1	1	0	3	3	2
HIBERNATE	0	0	0	0	0	0	1	0	1	0	2	2	2
EJB	0	0	0	0	0	0	1	1	0	0	2	2	2

Figure 4 Representative qualifications that belong to the intersection of IT market needs and university curricula

R2) The second cluster contains qualifications that are widely requested on the actual IT market, but that are poorly or not at all represented in the universities academic plans.

This cluster generally contains qualifications demanded by either emerging technologies or by enterprise-level activity (Figure 5).

In this case, the values of 1 in the matrix represent the monopole of an university over a qualification that is requested on the IT market. These qualifications can be used in the university's marketing strategy as a competitive advantage over the other educational institutions.

Qualification	Uni1u	Uni2u	Uni3u	Uni4u	Uni5u	T_u	Uni1m	Uni2m	Uni3m	Uni5m	T_m	T_edu	T_need
MICROSOFT SQL SERVER	0	0	0	0	0	0	0	0	0	0	0	0	7
AJAX	0	0	1	0	0	1	0	0	0	0	0	1	4
SHELL SCRIPTING	0	0	0	0	0	0	0	0	0	0	0	0	4
WINDOWS APPLICATIONS	1	0	0	0	0	1	0	0	0	0	0	1	3
APACHE	0	0	0	0	0	0	0	0	0	0	0	0	3
COM	0	0	0	0	0	0	0	0	0	0	0	0	3
JBOSS	0	0	0	0	0	0	0	0	0	0	0	0	3
WEBSPHERE	0	0	0	0	0	0	0	0	0	0	0	0	3
ASP	0	1	0	0	0	1	0	0	0	0	0	1	2
POSTGRESQL	0	0	0	0	1	1	0	0	0	0	0	1	2
DB2	0	0	0	0	1	1	0	0	0	0	0	1	2
ORACLE OLAP	0	0	0	0	0	0	1	0	0	0	1	1	2
MFC	0	0	0	0	0	0	0	0	0	0	0	0	2
SERVER ADMINISTRATION	0	0	0	0	0	0	0	0	0	0	0	0	2
WIN2003	0	0	0	0	0	0	0	0	0	0	0	0	2
WEB SERVERS	0	0	0	0	0	0	0	0	0	0	0	0	2
ATL	0	0	0	0	0	0	0	0	0	0	0	0	2
SILVERLIGHT	0	0	0	0	0	0	0	0	0	0	0	0	2
T-SQL	0	0	0	0	0	0	0	0	0	0	0	0	2
SHAREPOINT	0	0	0	0	0	0	0	0	0	0	0	0	2

Figure 5 Qualifications that are requested on IT market and that are little prepared in university formations

R3) The third cluster includes those qualifications that are prepared by universities in spite of the low demand on the actual IT market for those qualifications.

This cluster includes qualifications for research jobs and that are not applicable in an operational IT department (Figure 6). The number of ones in a formation column, if too large, can be considered evidence of a very theoretical curricula, with low practical applicability.

The second analysis intends to discover the percentage in which different combinations of undergraduate and master programs prepare demanded jobs on IT market. In order to accomplish this, for each job we get the set of qualifications needed as resulted from the web content mining process. From the educational field we consider the following possible formation tracks:

Qualification	Uni1u	Uni2u	Uni3u	Uni4u	Uni5u	T_u	Uni1m	Uni2m	Uni3m	Uni5m	T_m	T_edu	T_need
OPERATING SYSTEMS	1	1	1	1	1	5	0	0	0	0	0	5	16
ORACLE	1	1	1	1	1	5	0	0	0	0	0	5	11
JAVA	0	1	1	1	1	4	1	0	1	0	2	6	8
COMPUTER NETWORKS	0	1	1	1	1	4	1	0	0	0	1	5	8
MYSQL	0	0	0	1	1	2	0	0	0	0	0	2	7
JAVASCRIPT	0	0	1	0	1	2	0	0	0	0	0	2	7
C++	1	1	1	1	1	5	0	0	1	0	1	6	6
XML	1	1	1	0	1	4	0	0	1	0	1	5	6
PHP	0	1	1	1	1	4	0	0	0	0	0	4	6
OOP	1	1	1	0	1	4	0	0	0	0	0	4	6
TCP/IP	0	1	0	0	0	1	1	0	0	0	1	2	6
SECURITY	0	0	1	0	1	2	1	1	0	0	2	4	4
PL/SQL	0	1	1	0	0	2	1	0	0	1	2	4	4
ALGORITHMS AND DATA STRUCTURES	1	1	1	1	1	5	0	0	1	0	1	6	3
UML	0	1	1	0	1	3	1	0	1	1	3	6	3
ERP	1	1	0	1	0	3	1	0	0	0	1	4	3
DBMS	0	0	1	1	1	3	0	0	0	0	0	3	3
DATABASE DESIGN	1	1	1	0	0	3	0	0	0	0	0	3	3
VISUAL BASIC	0	0	1	1	0	2	0	0	0	0	0	2	3
ASP.NET	0	0	1	0	0	1	1	0	0	0	1	2	3
STRUTS	0	0	0	0	0	0	1	0	1	0	2	2	3
C#	1	0	1	0	1	3	0	0	0	0	0	3	2
DATABASE MANAGEMENT	0	0	1	0	0	1	1	1	1	0	3	4	2
JSP	0	0	1	0	0	1	1	1	1	0	3	4	2
ASSEMBLER	0	1	0	0	0	1	1	0	0	0	1	2	2
PROJECT MANAGEMENT	0	0	0	0	0	0	1	1	1	0	3	3	2
HIBERNATE	0	0	0	0	0	0	1	0	1	0	2	2	2
EJB	0	0	0	0	0	0	1	1	0	0	2	2	2

Figure 6 Qualifications that are prepared in university formations and are not demanded on IT market

Xu – undergraduate program in university X

XuYm - undergraduate program in university X followed by a master program in university Y, where X, Y $\in$ {Uni1, Uni2, Uni3, Uni4, Uni5}.

The resulting matrix contains float values obtained with the function (10).

g: J x US -> R,

$$g(j, f) = \begin{cases} \left(\sum_{q \in pr_{Job=j} JQ} M[q][f]\right) / \left(|pr_{Job=j} JQ|\right) *100 \text{ if f is of form Xu} \\ \left(\sum_{q \in pr_{Job=j} JQ} ((M[q][fu] \text{ AND } M[q][fm]) \text{ OR } (M[q][fu] \text{ XOR } M[q][fm])\right) / \left(|pr_{Job=j} JQ| \right) *100 \text{ if } f=fufm \text{ is of form XuYm} \end{cases} \quad (10)$$

Therefore, according to the matrix represented in table 1, for example, graduating the Uni1 undergraduate program (Uni1u) provides the fresh graduate with 50% of the qualifications needed in a „Deployment consultant" job, while following a track of Uni1 undergraduate program followed by a master in Uni1 in the IT field , provides the fresh graduate of this track (Uni1uUni1m) with 75% of the qualifications needed for the „Deployment consultant" job.

Jobs prepared by various university formation tracks and the corresponding percentages

Table 1

JOB	Uni1u	Uni2u	Uni3u	Uni4u	Uni5u	Uni1u Uni1m	Uni1u Uni2m	Uni1u Uni3m	Uni2u Uni1m	Uni2u Uni2m	Uni2u Uni3m	Uni3u Uni1m	Uni3u Uni2m	Uni3u Uni3m
C++ developper	28.6	39.3	35.7	28.6	32.1	53.6	32.14	32.14	50	42.86	42.86	53.57	35.71	35.71
Quality engineer	16.7	22.2	22.2	11.1	22.2	22.2	22.22	22.22	22.22	27.78	22.22	22.22	27.78	22.22
Web designer	14.3	7.14	21.4	21.4	28.6	28.6	14.29	14.29	21.43	7.14	7.14	28.57	21.43	21.43
.NET developper	21.1	15.8	21.1	5.26	21.1	26.3	21.05	21.05	21.05	15.79	15.79	26.32	21.05	21.05
Software engineer	10	40	30	20	20	30	10	20	40	40	40	30	30	30
IT consultant	37.5	37.5	25	12.5	25	50	62.5	62.5	50	62.5	50	37.5	62.5	50
Security analyst	22.2	66.7	77.8	55.6	66.7	66.7	33.33	22.22	77.78	77.78	66.67	100	77.78	77.78
Senior Developper	28.6	28.6	28.6	28.6	28.6	28.6	28.57	28.57	28.57	28.57	28.57	28.57	28.57	28.57
DB administrator	25	37.5	50	25	25	37.5	37.5	25	37.5	50	37.5	50	50	50
Deployment consultant	50	75	62.5	75	62.5	75	50	62.5	75	75	75	75	62.5	62.5
UNIX Administrator	12.5	18.8	18.8	18.8	18.8	18.8	12.5	12.5	18.75	18.75	18.75	18.75	18.75	18.75
Administrator IT Senior	40	60	60	40	60	100	60	40	80	80	60	80	60	60
analogue circuits engineer	0	50	0	0	0	0	0	0	50	50	50	0	0	0
SAP consultant	11.1	16.7	38.9	22.2	38.9	55.6	33.33	38.89	55.56	38.89	38.89	66.67	50	50
Network administrator	18.2	63.6	54.6	45.5	54.6	54.6	27.27	27.27	72.73	72.73	63.64	63.64	54.55	54.55
C# developper	50	50	50	25	25	75	75	75	75	75	75	75	75	75
Application designer	37.5	75	75	50	62.5	75	37.5	62.5	75	75	75	75	75	75
IT Helpdesk Support	25	50	75	50	75	75	50	25	75	75	50	75	75	75
Axapta Developper	50	62.5	62.5	25	50	62.5	50	62.5	62.5	62.5	62.5	75	62.5	62.5
ASP.NET developper	33.3	28.6	47.6	28.6	38.1	52.4	33.33	47.62	47.62	28.57	38.1	66.67	47.62	57.14
Web optimiser	0	25	16.7	16.7	25	8.33	0	0	25	25	25	25	16.67	16.67
Web developper	11.8	29.4	41.2	23.5	47.1	23.5	11.76	17.65	35.29	29.41	29.41	41.18	41.18	41.18
C++/Java developper	17.4	21.7	30.4	21.7	34.8	47.8	30.43	26.09	43.48	34.78	30.43	47.83	39.13	34.78
ORACLE Developper	0	33.3	50	33.3	33.3	33.3	0	50	33.33	33.33	66.67	50	50	83.33
System engineer	66.7	66.7	66.7	66.7	66.7	66.7	66.67	66.67	66.67	66.67	66.67	66.67	66.67	66.67
PHP Programmer	16.7	33.3	66.7	33.3	66.7	16.7	16.67	16.67	33.33	33.33	33.33	66.67	66.67	66.67

The Table 1 that we obtained represents a decision tool for all stakeholders: universities, HR managers, students.

For universities, it represents a snapshot of the real capabilities it can offer to its students and, combined with the granular analysis presented, it may provide some guidelines for universities of how to better adapt their curricula to the market needs.

For HR managers, it represents a valuable tool in the hiring decision making process. Given a set of job candidates that have no previous experience and are fresh graduates, HR personnel can score the candidates' compatibility with the advertised job according to Table 1.

For students, the table can represent a valuable career planning tool as it enlightens the formation tracks that the student should follow in order to obtain the best qualifications for a desired job.

Another important conclusion that can be raised refers to the necessity of a master program after the undergraduate studies in the Bologna system of education. As reflected in Table 1, a master increases with up to 40% the qualifications for a job, depending on the formation track.

4. Conclusions and Futures Directions of Studies

The strong points of the proposed framework include: an objective and automatic evaluation of the university formations compatibility with the market demand, the wide area of stakeholders that can obtain benefits by exploiting the results of this research, the uniqueness of this study in the Romanian IT sector until now.

A few weak points of the proposed model can be subject to further debates: the accurateness and objectivity of the educational plans description on university websites and the degree to which qualifications demanded for certain jobs on e-Recruitment websites are justified.

Considering the quality of data as being fair, the model provides a reliable decision model for all stakeholders implied: universities, company managers, students.

Based on these findings we aim to continue our research with a model of investments in IT human resources in small and medium enterprises. The research aims to evaluate different human resource strategies (training the employees, hiring fresh graduates, attract collaborators) in order to enrich the company's knowledge.

Acknowledgements

This paper represents a contribution of Ciolac Camelia Elena to the project „Doctoral Program and PhD Students in the education research and innovation triangle". This project is co funded by European Social Fund through The Sectorial Operational Program for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies. Also, this paper contains results of all authors carried out within the Project "Integrated platform for exploring the Collaborative Working Networks (CWN) based on a specific ontology structure oriented on manufacturing and services SMEs' needs", Code 92094/ 2008/ Partnerships.

References

1. Alexander, N.A. (2000). „The missing link: an econometric analysis on the impact of curriculum standards on student achievement". *Economics of Education Review*, Volume 19, Issue 4, 351-361
2. Burgoyne, J.G. (1975). „Towards an evaluation of a business school graduate programme". *Omega*, Volume 3, Issue 4, 475-482
3. CNPFA (nd). „Occupational Standards". Retrieved from http://so.cnfpa.ro/so/
4. Cyrenne, P., & Grant, H. (2009). „University decision making and prestige: An empirical study". *Economics of Education Review*, Volume 28, Issue 2, 237-248
5. Forgionne, G.A. (1992). „Streamlining information science research and education". *Journal of Information & Management*, Volume 23, Issue 3, 141-147
6. Johnes, J. (2006). „Measuring teaching efficiency in higher education: An application of data envelopment analysis to economics graduates from UK

Universities 1993”. *European Journal of Operational Research*, Volume 174, Issue 1, 443-456

7. Kitchenham, B., Budgen, D., Brereton, P., & Woodall, P. (2005). „An investigation of software engineering curricula”. *Journal of Systems and Software*, Volume 74, Issue 3, 325-335
8. Lee, S., Koh, S., Yen, D., & Tang, H. L. (2002). „Perception gaps between IS academics and IS practitioners: an exploratory study”. *Journal of Information & Management*, Volume 40, Issue 1, 51-61
9. Li, E.Y., Yen, H.J.R., & Cheng, C.Y.J. (2008). **„**A fit-gap analysis of e-business curricula and job demand in Taiwan and the US”. *Journal of Computers & Education*, Volume 51, Issue 3, 969-987
10. Luban, F. (2005). „Lifelong Learning and Knowledge Management”. *Economia. Seria Management*. Volume 8, Issue 1, 98-117
11. Nicolescu, O., (2009).”Multidimensional and Multicausal Approach of ManagerialChange inside Organisations”. *Review of International Comparative Management*, Vol. 10, Issue 5, 853-863
12. Ono, H. (2008). „Training the nation’s elites: National–private sector differences in Japanese university education”. *Research in Social Stratification and Mobility*, Volume 26, Issue 4, 341-356
13. Ratiu-Suciu, C., Stoica, M.D. (2004). „Advances in Modelling the Phenomena Occuring on the Market of Knowledge as Concerns the Investment”. *Proceedings of XI SIGEF Congress Reggio-Calabria – Messina Italy – Techniques and Methodologies for the Information and Knowledge Economy*. ISBN 88-8296-146-X, Fazea Editore, 210-221
14. Reeves, G.R., Bussom, R.S. (1980). „Information systems curricula: Business versus academic viewpoints”. *Journal of Information & Management*, Volume 3, Issue 6, 227-235
15. Richards, M., Pelley, L. (1994). „The ten most valuable components of an information systems education”. *Journal of Information & Management*, Volume 27, Issue 1, 59-68
16. Salas-Velasco, M. (2009). „Beyond lectures and tutorials: Formal on-the-job training received by young European university graduates”. *Research in Economics*, Volume 63, Issue 3, 200-211
17. Wenger, J. (2003). „The future of programming languages: evidence to support a midwest university information systems curriculum”**.** *Journal of Systems and Software*, Volume 65, Issue 2, 163-167