A Smart Approach for Matching, Learning and Querying Information from the Human Resources Domain⋆

Jorge Martinez-Gil, Alejandra Lorena Paoletti, and Klaus-Dieter Schewe

Software Competence Center Hagenberg GmbH
Softwarepark 21, 4232 Hagenberg, Austria
{jorge.martinez-gil,lorena.paoletti,kd.schewe}@scch.at
http://www.scch.at

Abstract. We face the complex problem of timely, accurate and mutually satisfactory mediation between job offers and suitable applicant profiles by means of semantic processing techniques. In fact, this problem has become a major challenge for all public and private recruitment agencies around the world as well as for employers and job seekers. It is widely agreed that smart algorithms for automatically matching, learning, and querying job offers and candidate profiles will provide a key technology of high importance and impact and will help to counter the lack of skilled labor and/or appropriate job positions for unemployed people. Additionally, such a framework can support global matching aiming at finding an optimal allocation of job seekers to available jobs, which is relevant for independent employment agencies, e.g. in order to reduce unemployment.

Keywords: e-Recruitment, Knowledge Engineering, Knowledge-based Technology

1 Introduction

Some of the major problems concerning the labor market are the complicated situation of the job market in many countries around the world and the increased geographical flexibility of employees. This situation makes companies to often receive a huge number of applications for every open position. Therefore, the costs of manually selecting potential candidates is usually high. For this reason, most companies would like to decrease the costs when publishing job postings

and selecting appropriate applicants from such a plethora of potential candidates [1]. It is also important to remark that unsuccessful job applicants often complain on the lack of transparency in the recruitment processes, and they often wish to receive detailed arguments, or at least, some information about the strengths and flaws of their profiles [29]. However, they do not receive any kind of feedback very often since this has to be done manually by the other part, and it is quite expensive, in terms of time and resource consumption, for the companies to do that [18].

This complicated situation leads us to think the accurate matching of curriculum vitae (CV) and job offers is very important for employers and job seekers. Therefore, the development of computational methods to optimize the recruitment processes should be of high importance in our current society [15]. Furthermore, such an approach could be beneficial for public and private employment agencies which could perform an analysis to determine the most needed qualification and training courses that would improve the skills of job seekers with respect to the market demands. As a result, a higher occupation rate could be achieved [23].

Currently, existing software solutions in this field are based on syntactic matching, i.e. for a requested profile, existing solutions check how many of the requested terms are overlapped in the candidate profile [22]. This fact ignores similarity between skills, e.g. programming skills in C++ or Java would be rated similar by a human expert [8]. Improving this primitive form of matching requires at least taking hierarchical dependencies between education or skill terms into account. To do that, various taxonomies have already been developed such as DISCO competences[1], ISCO[2] and ISCED[3]. These taxonomies play a central role in our research, since we can exploit them for achieving a more realistic mediation between open employment offers and suitable candidates. Therefore, our major contribution can be summarized as follows:

– We propose here a novel approach for the automatic matching, learning and efficient querying of information from the Human Resources (HR) domain. This approach is based on new methods that appropriately handle traditional limitations, including the uncertainty of human language, the incapability to exploit background knowledge, and the lack of a truly semantic mediation. Additionally, this approach could be of great interest for education and training institutions which could perform analysis to determine the most needed skill sets that would improve the skills of job seekers with respect to the available positions.

The rest of this paper is organized as follows: Section 2 describes the state-of-the-art concerning realistic matching, learning and querying information concerning HR. Section 3 describes the matching problem we are facing and why

---

[1] http://www.disco-tools.eu
[2] http://www.ilo.org/public/english/bureau/stat/isco/isco08/index.htm
[3] http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx

it is relevant in this context. Section 4 explains how the HR field could be benefit from a framework for learning to rank candidates. Section 5 discusses our approach's capability for querying, and finally, we draw conclusions and put forward future lines of research.

2 State-of-the-art

The problem of automatically matching job offers and applicant profiles has been studied in the scientific literature [2], but the complex nature of the problem we have to face, which involves the use of free text by employers (when writing their job offers) and by employees (when writing their application), makes developed solutions in this context unable to reach a high degree of success [18]. Some works have offered partial solutions based on the use of controlled vocabularies (i.e. ontologies) in order to fairly alleviate some problems concerning semantic heterogeneity [5] but there are still some key challenges that should be addressed [24].

One of these most important challenges is that the process of matching CVs and job offers is usually done without use of any knowledge base (KB). Instead, overlapping information is computed. In fact, according to the researched literature, a wide range of solutions for job and profiles matching have been addressed by a variety of techniques, ranging from simple bipartite graph matching [7], to vector based techniques taken from classical information retrieval [6], to record matching in databases [30].

Algorithms for bipartite graph matching try to find optimal solutions when trying to maximize the number of matching relation. However, these approaches rely on assigning costs to every match between curriculum and profiles. When the costs are assigned manually, knowledge about them is completely subjective, and therefore it becomes very difficult to revise [3]. Moreover, an approach maximizing the number of matches may provide a bad service to users: for example, person P1 could have the best match for job profile J1, but she might be suggested to take job J2 just because J1 is the only available job for person P2 [13]. This means that from a strictly user-centric viewpoint, maximizing the number of matches is not the feature that could face our problem.

More sophisticated approaches are based on database techniques for record matching [12] or information retrieval [21]: feature vectors, analytical geometric similarity, weighted criteria, keyword-based search, assessment based on recall and precision [17]. In case of non-suitable highly ranked profiles human expertise can be used to correct inaccuracies. The problem with these techniques is that they are not suited for dealing with incomplete information usually present in scenarios of this kind. In fact, information about profiles is not always complete, not only because some information is unavailable, but also because some details are considered irrelevant by either the employer or the applicant. Trying to force to use an interface for entering profiles with long and tedious forms to be filled in, is the most often adopted solutions to this problem [27].

Among the problems concerning learning, the task of learning to rank has probably received the most attention in the machine learning literature in recent years. In fact, a number of different ranking problems have been introduced so far. The ranking module is one of the most important modules in a Human Resources Management (HRM) system. For a given job offer there may be hundreds or thousands of relative candidates but only a few of them are to be shown to the expert at a time. Therefore, it is very important to fetch the most relevant candidates and display them to the expert. This means that the way that top candidates are presented decide the success of the HRM system, and therefore, each one of the entries is important.

The major challenge here is to use the expert behavior as a feedback. However, some researchers are skeptical about using this kind behavioral data as a feedback because there are various biases involved in taking behavior into consideration. They show that there exists some presentation bias, which is the bias involved when experts instinctively prefers some candidates in relation to others. It means some candidates are more likely to get better attention from experts and other candidates are not given the proper attention even though they are more relevant. However, it is possible to find useful strategies to solve this bias [11].

In practice, when proposing solutions concerning ranking, we think it is a good idea to consider the algorithm Okapi BM25 [25] as the baseline to compare new approaches in this field. The reason to choose an algorithm of this kind is that it is widely used by software systems to rank matching candidates according to their relevance to a given search offer. Okapi BM25 is considered the state-of-the-art among the methods using a syntactic approach [14]. Therefore, any new method in the field of automatic matching should prove its effectiveness when compared to it.

With respect to querying knowledge bases, in particular in the HR domain, the commonly investigated approach is to find the best k (with $k = 1$ in most cases) matches for a given profile (applicant profile or job offer) [4]. Though this constitutes what is commonly known as top-k-queries, a systematic investigation of such kind of queries is still missing. Top-k-queries have been thoroughly investigated in the field of databases, usually in the context of the relational data model [10], but the study of such queries in the context of knowledge bases has not yet been done. The expectation is of course, that many of the results in the relational data model can be easily adopted to this case. In particular, the focus on a single relation, i.e. the matching, as the driver for the querying, is expected to ease the extension.

In addition to top-k-queries the interest in partial orders in extended matching relations leading to skyline queries as well as global matching optimization and gap analysis place further challenges on matching-related querying of knowledge bases that have not yet been investigated. The classification of most relevant types of queries and the adaptation of corresponding state-of-the-art approaches in databases should be the emphasis in the future. The expected results are supposed to support the efficient answering of such queries.

3   Matching Information from the Human Resources Domain

In this context, semantic matching is a well know problem whereby two entities in a knowledge base are assigned a score based on the likeness of their meaning [16]. Automatically performing semantic matching is considered to be one of the pillars for many computer related fields since a wide variety of techniques rely on a good performance when determining the meaning of data they work with [19].

More formally, we can define semantic matching as a function $\mu_1 \text{ x } \mu_2 \rightarrow \text{R}$ that associates the degree of correspondence for the entities μ_1 and μ_2 to a score $s \in \text{R}$ in the range $[0, 1]$, where a score of 0 states for not correspondence at all, and 1 for total correspondence of the entities μ_1 and μ_2.

Traditionally, the way to compute the degree of correspondence between entities has been addressed from two different perspectives: using semantic similarity measures and semantic relatedness measures. Fortunately, recent works have clearly defined the scope of each of them. Firstly, semantic similarity is used when determining the taxonomic proximity between entities. For example, automobile and car are similar because the relation between both terms can be defined by means of a taxonomic relation. Secondly, the more general concept of semantic relatedness considers taxonomic and relational proximity. For example, nurse and hospital are not completely similar, but there is still possible to define a naive relation between them because both belong to the world of healthcare [19].

In most of cases, the problem to face is much more complex since it does not only involve the matching of two individual entities, but two complete documents (applicant profile or job offer). This can be achieved by computing a set of semantic correspondences between individual entities belonging to each of the two documents. A set of semantic correspondences between entities is often called an alignment. It is possible to define formally an alignment A as a set of tuples in the form $\{(id,\ \mu_1,\ \mu_2,\ r,\ s)\}$, where id is an unique identifier for the correspondence, μ_1 and μ_2 are the entities to be compared, r is the kind of relation between them, and s the score in the range $[0, 1]$ stating the degree of correspondence for the relation r.

Therefore, when matching two documents, the challenge that scientists try to address consists of finding an appropriate semantic matching function leading to a high quality alignment between these two knowledge bases. Quality here is measured by means of a function $A \times A_{ideal} \rightarrow R \times R$ that associates an alignment A and an ideal alignment A_{ideal} to two real numbers $\in [0, 1]$ stating the precision and recall of A in relation to A_{ideal}.

Precision represents the notion of accuracy, that it is to say, states the fraction of retrieved correspondences that are relevant for the matching task (0 stands for no relevant correspondences, and 1 for all correspondences are relevant). Meanwhile, *recall* represents the notion of completeness, thus, the fraction of relevant correspondences that were retrieved (0 stands for not retrieved correspondences, and 1 for all relevant correspondences were retrieved).

Applying this kind of techniques fits well in the HR scenario. The reason is that these techniques can be used for going beyond the literal lexical match of words. In this way, when analyzing the curriculum of job candidates, this kind of techniques can operate at the conceptual level when comparing specific terms (e.g., Finance) also yields matches on related terms (e.g., Economics, Economic Affairs, Financial Affairs, etc.). As another example, in the healthcare field, an expert on the treatment of cancer could also be considered as an expert on oncology, lymphoma or tumor treatment, etc [9]. The potential of this kind of techniques is that it can support Human Resource Management when leading to a more quickly and easily cut through massive volumes of potential candidate information, but without giving up the way human experts take decisions in the real world.

4   Learning Information from the Human Resources Domain

The problem of learning can be defined as given a pair of objects (jo, ap_i) together with a measure of their suitability $y_i \in \mathbb{R}$. The goal is to learn a function $f(jo, ap_i) \approx y_i$ that approximates for every new labeled triplet example (jo, ap_i, y_i), where jo is a job offer, ap_i is a list of applicant profiles, and y_i is the associated list of scores of each ap_i for the job offer jo.

After many discussions with professionals from the Human Resources sector, we agreed this challenge has not an unique solution. The reason is that every HR professional evaluating different cases could propose different results. This makes us thinking that we should work towards an adaptive approach by means of automatic matching learning. This approach should be able to calculate the transformation cost of a given profile into a requested job offer, so that profiles with higher transformation cost should rank worse than those with lower cost. In this way, our approach should be able to replicate the results from the human experts. This means that for each person aiming to use a solution of this kind, we should train a model for capturing its know-how or preferences by means of an initial training stage. Thinking on a model of this kind is far from being trivial. However, we assume that a generic solution for this problem should be characterized by the following core attributes: a) a base distance between sets, b) some background knowledge to compute the replacement cost, c) the desired cost of insertion and deletion of new elements, d) the way to weight elements, either a multiplicative or an additive preference

Please note that if we work with different relevant subsets (education, skills, languages, etc.) the transformations costs could be different for each subset, so the final cost should be an aggregation of the partial costs for each segmented group. Once we get a solution, the way to determine if this solution is satisfactory could be defined as the correlation between this achieved solution and an ideal one.

Concerning a), we can formally define our distance between two sets as the minimum number of single-elements edits (i.e. insertions, deletions or substi-

tutions) required to change one set into the other. It is very appropriate for computing the transformation costs from a CV into a job offer.

Concerning b), setting up adequate knowledge bases that capture recruitment terminology in a precise and easily extendable way is a crucial success factor. So far, no such knowledge bases exist. However, our existing matching technology is based on valuable recruitment taxonomies. These taxonomies are structured thesaurus and vocabularies for the description of skills in different scenarios such as the education, job market and training courses respectively. These taxonomies provide us a complete skill and competence classification which is based on existing European and international standards and classifications, and therefore, represent a terminological basis for the standard description of skills, competences, occupations as well as applicant profiles, job vacancies, and job requirements, etc. or for describing professional degrees, study programs, courses, and so on. To illustrate why taxonomies are important for us, let us suppose that a job offer requests a person skilled in Java, and we have a candidate who is skilled in JavaScript. We can compute the shortest path between Java and JavaScript in the recruitment taxonomy. The transformation cost can be based on the length of this path. In this way, short paths leads to low replacement costs, and on the contrary; longer paths may lead to higher replacement costs. If there is no path between them, or even this path is not appropriate enough (i.e. too long) then we can consider insertion and deletion costs.

Concerning c), Suitability of an applicant profile ap_i to a job offer jo needs also to consider the minimum cost of element insertions and deletions which transforms the applicant profile ap_i into the job offer jo. These costs are going to be used when an applicant profile have a different number of elements than those requested in the job offer or computing the replacement cost between elements is not possible. The computation of these costs is of vital importance because it helps us to characterize the behavior of the people who was involved in the training stage. Insertion cost is an estimation of how much it could cost to a potential candidate to acquire an element requested by the job offer.

Deletion cost is an estimation about the impact of having a not requested element. For example, an expert could think that candidates holding not requested elements could be unhappy, unmotivated, could request a higher salary or be willing to leave the company in a short period of time. The penalty to be applied can be high, if the person in the training phase tends to penalize overqualification, null if the person does not care about additional (although not requested) elements, or even negative, if the person training the model thinks that additional elements are far from hurting. It is also important to note that we cannot have an unique value for insertions and deletions costs. For instance, it is much more expensive (in terms of effort, time and money) acquiring a new university degree that some certain level of mastery in a programming language or technology.

Concerning d) the weighting schema is the way a person could increase or decrease the importance of the elements within a given set. Considering a weighting schema is important because it allows job recruiters giving more importance to

some facts like years of experience, level of mastery or simply stating priorities for filling a position.

4.1   Querying Information from the Human Resources Domain

One of the main requirements from the HR application domain leads to queries on a knowledge base of job offers and candidate CVs. Ignoring the inherent inferential capability given by knowledge bases. Each knowledge base is also a database in the sense that there is a schema, i.e. the concepts and roles in the TBox, and a set of instances, i.e. the ABox. Therefore, adopting database technology as key method to address the querying problems is a natural idea [23].

In database technology effective and efficient query processing is a core area with a tradition since decades. Recently, two classes of queries, top-k-queries and skyline queries have attracted the interest of researchers [26]. For top-k-queries assume that a query q produces an answer set $\mathcal{A}$ that is totally ordered. Then a query top-$k(q)$ will select the k largest elements of $\mathcal{A}$ as the answer. While performing a sorting operation and a cut-off of the largest k elements are straightforward in theory, the key problem with top-k-queries is efficiency on very large databases, for which supporting data structures and rewriting techniques that enable the computation of the k largest answers without computing first all answers. Similarly, skyline queries ask for all maximal elements in an answer set $\mathcal{A}$ to a query q, where $\mathcal{A}$ is assumed to be partially ordered.

We think that top-k- and skyline queries are essential for the core of matching related queries, where the (partial) order is defined by the matching measures. In case of simultaneous use of several matching measures a partial order may result. Therefore, the key research question is to adopt the solutions from database technology to the area of knowledge bases, which boils down to investigating efficient storage of the ABox including matching measures. For the data structures supporting the subsumption hierarchy it is envisioned that rings and spiders [20] known from network databases and revived in object-oriented databases can be adopted. These structures are known for excellent performance in support of queries that exploit hierarchical data structuring. Furthermore, indices based on partial fractions may also be exploited for this purpose [28].

It is further anticipated that skyline queries will also play an important role for gap analysis, which should result in minimally enlarged filters that guarantee improved matching results. That is, we have to exploit a partial order on filters for such queries. The enlargement itself requires for data structures supporting neighborhoods, which will be a new notion that has to be defined and for which suitable storage representations have to be found. With such extensions it should be possible to exploit state-of-the-art techniques for skyline queries to support the application needs. Furthermore, specific query optimization techniques will be needed.

The adaptation and extension of query optimization is also the method that is needed to support global matching with respect to some optimization criteria. As the optimization criteria will lead again to a partial order, this gives another

class of skyline queries, so the remaining problem is efficiency, for which query optimization is due.

5   Discussion

We think our approach lead to a number of qualitative advantages over the state-of-the-art in this field. These advantages are in the direction of those mentioned in [18]. In fact, we can summarized them in the following four major points:

1. Our approach for realistic matching learning can help players from the HR industry to go beyond syntactical matching of job offers and applicant profiles. This represents a great advantage over the current state-of-the-art since our approach tries to give more opportunities to the good job candidates, but also allows job recruiters to identify potential talent which otherwise may remain blurred among such a plethora of applicants profiles.
2. Our approach can help to eliminate the need for job recruiters to have deep and specialized knowledge within an industry. This is mainly due to this approach is able to model knowledge from a lot of industrial domains. Then this knowledge can be used as a support when performing matching process so that the results can be very similar to those produced by an expert from that field.
3. Our approach can provide feedback to the applicants that did not get the job. The matching process is traceable and this means that some interesting reports can be automatically delivered to the applicants. These reports can help these applicants to determine the reasons they were not selected for the job position as well as to assess their strengths and weaknesses when applying for similar jobs in the future.
4. Our approach allows to leveling the odds for those job applicants with less ability when preparing their resumes. The reason is that an algorithm will perform the matching process automatically. The result from this process is independent of the way the curriculum is presented. Therefore, this technique helps to promote equal opportunities.

6   Conclusions

We have presented a novel approach for the timely, accurate and mutually satisfactory mediation between open employment offers and suitable candidates. The rationale behind this research approach is to facilitate public and private recruitment agencies as well as employers and job seekers around the world to reduce the costs and time to find relevant matches between job offers and applicant profiles.

The major conclusion we can extract is that an approach of such kind may be able overcome the traditional limitations in this field. 1) Concerning the uncertainty when dealing with natural language: our solution forces to describe either job offers and applicant profiles using a common vocabulary. This fact avoid

the problem of dealing with different representations of a same concept. 2) Concerning the incapacity of current approaches to exploit background knowledge: our approach is able to get benefit from external knowledge bases that support our work within the process of making estimations about the cost of acquiring a new skill or competence. 3) Concerning achieving a truly realistic mediation: our approach is able to go beyond syntactic matching of job offers and applicant profiles. Additionally, some practical impact can be envisioned, e.g. capability to reduce unemployment.

As a future work, we plan to develop a benchmark data set in order to assess the quality of the new matching proposals based on this and other different paradigms. The importance of this benchmark data set is high since researchers need some data to test their solutions. One solution could consists of automatically generated random data sets. The problem is that randomly generated data sets are not useful since they do not have realistic distributions, and it is not always easy to understand whether the achieved results are meaningful and correct. Real or quasi-real data is much better in this particular application scenario. For these reasons, we are working towards a benchmark data set that reflects the way real HR experts take decisions in real scenarios and allows research community testing and training new approaches under truly realistic conditions.

References

1. Bizer, C., Heese, R., Mochol, M., Oldakowski, R., Tolksdorf, R., Eckstein, R. The Impact of Semantic Web Technologies on Job Recruitment Processes. Wirtschaftsinformatik 2005: 1367-1382.
2. Bradley, K., Smyth, B. Personalized information ordering: a case study in online recruitment. Knowl.-Based Syst. 16(5-6): 269-275 (2003).
3. Cali, A., Calvanese, D., Colucci, S., Di Noia, T., Donini, F.M. A Logic-Based Approach for Matching User Profiles. KES 2004: 187-195.
4. Chakrabarti, K., Ortega-Binderberger, M., Mehrotra, S., Porkaew, K. Evaluating Refined Queries in Top-k Retrieval Systems. IEEE Trans. Knowl. Data Eng. 16(2): 256-270 (2004).
5. Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F.M., Mongiello, M., Mottola, M. A Formal Approach to Ontology-Based Semantic Match of Skills Descriptions. J. UCS 9(12): 1437-1454 (2003).
6. Faliagka, E., Tsakalidis, A.K., Tzimas, G. An Integrated E-Recruitment System for Automated Personality Mining and Applicant Ranking. Internet Research 22(5): 551-568 (2012).
7. Farber, F., Weitzel, T., Keim, T. An Automated Recommendation Approach to Selection in Personnel Recruitment. AMCIS 2003: 302.
8. Garcia Sanchez, F., Martinez-Bejar, R., Contreras, L., Fernandez-Breis, J.T., Castellanos Nieves, D. An ontology-based intelligent system for recruitment. Expert Syst. Appl. 31(2): 248-263 (2006).
9. Chaves-Gonzalez, J.M., Martinez-Gil, J. Evolutionary algorithm based on different semantic similarity functions for synonym recognition in the biomedical domain. Knowl.-Based Syst. 37: 62-69 (2013)

10. Ilyas, I. F., Aref, W. G., Elmagarmid, A. K.. Supporting top-k join queries in relational databases. VLDB J. 13(3): 207-221 (2004).
11. Joachims, T. The Value of User Feedback. ECIR 2011: 6.
12. Kessler, R., Bechet, N., Roche, M., Torres-Moreno, J.M., El-Beze, M. A hybrid approach to managing job offers and candidates. Inf. Process. Manage. 48(6): 1124-1135 (2012).
13. Kuokka, D., Harada, L. Integrating Information via Matchmaking. J. Intell. Inf. Syst. 6(2/3): 261-279 (1996).
14. Lv, U., Zhai, C. A Log-Logistic Model-Based Interpretation of TF Normalization of BM25. ECIR 2012: 244-255.
15. Malinowski, J., Keim, T., Wendt, O., Weitzel, T. Matching People and Jobs: A Bilateral Recommendation Approach. HICSS 2006.
16. Martinez-Gil, J., Aldana-Montes, J.F. Reverse ontology matching. SIGMOD Record 39(4): 5-11 (2010).
17. Martinez-Gil, J., Aldana-Montes, J.F. Evaluation of two heuristic approaches to solve the ontology meta-matching problem. Knowl. Inf. Syst. 26(2): 225-247 (2011).
18. Martinez-Gil, J. An Overview of Knowledge Management Techniques for e-Recruitment. JIKM 13(2) (2014).
19. Martinez-Gil, J. Automated knowledge base management: A survey. Computer Science Review 18: 1-9 (2015).
20. Mylopoulos, J., Brodie, M.L. Knowledge Bases and Databases: Current Trends and Future Directions. IS/KI 1990: 153-180.
21. Meo, P., Quattrone, G., Terracina, G., Ursino, D. An XML-Based Multiagent System for Supporting Online Recruitment Services. IEEE Transactions on Systems, Man, and Cybernetics, Part A 37(4): 464-480 (2007).
22. Mochol, M., Wache, H., Nixon, L.J.B. Improving the Accuracy of Job Search with Semantic Techniques. BIS 2007: 301-313.
23. Paoletti, A.L., Martinez-Gil, J., Schewe, K-D: Extending Knowledge-Based Profile Matching in the Human Resources Domain. DEXA (2) 2015: 21-35.
24. Racz, G., Sali, A., Schewe, K-D. Semantic Matching Strategies for Job Recruitment: A Comparison of New and Known Approaches. FoIKS 2016: 149-168.
25. Robertson, S. E., Walker, S., Hancock-Beaulieu, M. Experimentation as a way of life: Okapi at TREC. Inf. Process. Manage. 36(1): 95-108 (2000).
26. Soliman, M. A., Ilyas, I. F., Ben-David, S. Supporting ranking queries on uncertain and incomplete data. VLDB J. 19(4): 477-501 (2010).
27. Straccia, U., Tinelli, E., Colucci, S., Di Noia, T., Di Sciascio, E. A System for Retrieving Top-k Candidates to Job Positions. Description Logics 2009.
28. Theobald, M., Weikum, G., Schenkel, R. Top-k Query Evaluation with Probabilistic Guarantees. VLDB 2004: 648-659.
29. Thielsch, M. T., Traumer, L., Pytlik, L. E-recruiting and fairness: the applicant's point of view. Information Technology and Management (ITM) 13(2):59-67 (2012).
30. Tinelli, E., Cascone, A., Ruta, M., Di Noia, T., Di Sciascio, E., Donini, F.M. I.M.P.A.K.T.: An Innovative Semantic-based Skill Management System Exploiting Standard SQL. ICEIS (2) 2009: 224-229.