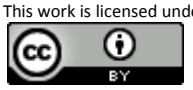
Available online at: https://ijact.in

Date of Submission	21/10/2018
Date of Acceptance	28/11/2018
Date of Publication	28/12/2018
Page numbers	2962-2968 (7 Pages)

This work is licensed under Creative Commons Attribution 4.0 International License.



ISSN:2320-0790

An International Journal of Advanced Computer Technology

ONTOLOGICAL APPROACH FOR COMPETENCE MANAGEMENT USING K-NN WITH R

Mohamed Fakir, Bouzekri Moustaid

Computer Sciences Department, Sultan Moulay Slimane University, Beni-Mellal, Morocco
fakfad@yahoo.fr, boumousta@gmail.com

Abstract: With the advent of the movement of the immaterial economy which leads us towards an economy and a company of knowledge, the importance is given to the concepts of knowledge and competencies for better immaterial capital development. Consequently, the management of the 'human capital' became both a necessary and a complex matter for the Human Resources professionals who are concerned about employee's intellectual capitals (knowledge management) all along with their competences.

The management of competence is a topic that has always interested in several research teams and was the object of several publications of scientific work. The research of qualified staff is one of the most important axes of this work. It lies in the scope of the decision-making aid for the selection of one person starting from a group of candidates in order to achieve the goals of employment with well defined characteristics.

In this context, we present various approaches referring to a competence, to its definitions, its characterizations and its methods of treatment. Our first contribution will be directed towards the proposal of a typology and a characterization of a job and competences and a system of evaluation of the level of these competences.

The semantic web technologies and those of the data mining are chosen for the implementation of this contribution. Indeed, competence is a field of knowledge with excellence. Thus, the field of the semantic web offers panoply of standards to represent knowledge and to handle them. The whole of this knowledge will be represented in the form of ontology that will be treated with the external resource. WordNet, in order to guide the research of candidates with the qualified skills to occupy a given job. The k-nearest neighbor algorithm, or quite simply kNN and the languages R and SPARQL are used for this implementation.

Keywords: competences, knowledge, ontology, KNN, R Language, SPARQL, HR-XML, HRIS

I. INTRODUCTION

The effectiveness of the human capital was considered before as a factor which does not have interest in the creation of the wealths of the countries. Competences and the human knowledge were ignored and underestimated compared to the physical effort in the creation of wealths. However, since the beginning of the 18th century, the development of competences is perceived like a crucial key of productivity and competitiveness in the public sector as well as in the private sector. The interest of learning as well as the competence and knowledge management is a striking fact of the current decade.

In this context, several initiatives were taken. For example, in France, two recent stages concretized, the importance of the concept of competence:

The skills assessment created in 1991:

It is a law of Programming for the Social cohesion of January 18th, 2005 imposing the approach of the management planning of employment and competences (GPEC)

Morocco, with the framework of its programme of reform of the public administration is committed beyond 2008, in the improvement of the effectiveness of the State, thanks to the introduction of a new management system of human resources. The installation of the Management planning of headcount, Employment and Competences (GPEEC) in the great ministries constitutes one of the most important pillars in this reform [1].

The National Initiative for Human development (INDH) is also an innovative initiative in Morocco for social and human development. Since in the Eighties the concept of competence is used more and more in various fields and disciplines: cognitive psychology, sociology, sciences of management, sciences of data processing and the artificial intelligence etc. but, it is not yet well defined. Actually, what poses a blur around is its definitions and the methodologies of its use makes it consequently a true problem for its treatment by the computer systems and in particular the human resources information systems (HRIS). The competence is therefore a privileged research topic. The first research into the problem of assignment is by formal approaches emanate primarily from the field of scheduling techniques [2]. Recently, fuzzy logic is the most used in these works. It is sometimes combined with other tools. With its appearance at the beginning of 1990, the concept of ontology has quickly become essential as a way to explain explicitly the objects of a domain for its shared understanding. Ontology occupies thus an important role in the process of competence and knowledge management. Our contribution uses an ontological approach for the management of competences by proposing a model of characterization of competences and an approach to exploit this ontology coupled with the external resource WordNet. The object is to lead the research for the candidates who have the necessary skills for a given job. We have chosen the algorithm of k-nearest neighbor. This article was structured in three chapters, an introduction, the material and method and finally the results and discussion.

I. THE MATERIAL AND METHOD

A. Definition of Competence:

The concept of competence is the object of many definitions through various scientific opinions which study it. This generates a semantic ambiguity and thus a problem of appropriation of this concept by the organizations which seek stable elements on which they can base their activities and their decisions. It is thus important to present the principal definitions of competence in order to be able to draw meaningful conclusions which will be useful thereafter in our contribution. We distinguish two different approaches relating to the definition of competence either as a set of resources or like a way of acting.

In the literature, it is fairly customary to find definitions that compare competence to a process. [3] defines the competence as "the production process of social performance, economic performance etc. in a concretely delimited context".

It is a generating process of the end product which is the performance. Boterf [4] defines a competent person as "a person who is able to mobilize, in an effective way his/her various resources such as operations of reasoning, knowledge, activations of the memory, evaluations, relational capacities or behavioral diagrams in diversified contexts or increasingly complex tasks.

Competence is a set of resources:

Competence is defined as "a sum of knowledge, know-how and of knowing to be". This definition is mostly used and is adopted in the regulation & governing of vocational training and in multiple projects as the case in Morocco [1] .

The know-how is a technical skills related to the act of realization of the tasks. The behavioral competences are related to the behaviors of the collaborator in his work environment. It is the knowledge-being which characterizes each individual. In the Anglo-Saxon literature, the definition of competence rests on work of [5]. Knowledge and know-how are described as "hardware competences", the other categories relating to the attitudes, behaviors, milked personality and motivations are gathered under the category "software competences".

The technical skills are thus generally known or easily identifiable in the context of the work of the collaborator. The components of the behavioral competence pose problems of identification and evaluation,

B. Characterization of competences and definition of a framework:

Various definitions which we presented cannot be exploited directly to build out a computer system. The characterizations of competences by measurable components by quantitative or qualitative measurements are essential. This characterization makes it possible to define a typology of the components and a rating scale of the levels.
Several terms were proposed in the literature to indicate components of the competence [6] [7].
For the evaluation of the level of competences, the taxonomy of the teaching objectives of [8] is one of the oldest references of evaluation of resources. Other scales of measurements are proposed [9], [10], [1].

C. Different approaches from treatment of competence

From the bibliographical study, we structured most of the scientific work according to the following axes:

- Identification and allowance of competences
- human resource information system HRIS,
- Tools specialized in the management of competences
- Standardization HR-XML:

C.A.1 The identification and the allowance of competences:

The identification of competences is conceived as a process of discovering the competences that the employees have at a given moment. It is a fundamental stage where the decisions depend on management of competences at all the horizons. The allowance of competences and the team building relate to the human resource allocation to the missions which correspond best to their competences and the building of a team work performance.

The research which integrate the problems of assignment by formal approaches emanate primarily from the field of scheduling. However, this work considers the human resources in the same way as a material resource that they apprehended only in terms of availability and capacities [11].

This was covered recently by the scientific research which proposes methods based primarily on fuzzy logic, hinges on the calculation of the distances between the competences required by the missions of the team and those which are acquired by the actors candidates [12], [13].

The method of fuzzy logic is sometimes combined with other tools as it is the case for the method suggested [Fitzpatrick &al., 2005] that combines a heuristic approach and a mathematical method.

The recourse to the algorithms for pairing between required/acquired competences constitutes the scientific research work which fit in this axis [14] and [15].

C.A.2 Tools specialized in the management of competences*:*

The computer tools were created to assist the managers in their mission of management of competences. An example of computer system of assistance to the management of competences tool is SEE-K (tool of analysis and visualization of information) of company TRIVIUM. It is software based on the works of Authier and Pierre Levy using the theories of the trees of knowledge [16].

One can find other systems which were developed based on models suggested for the structuring of competences [17], [7].

C.A.3 *Human Resources information system HRIS:*

The integration of the concept of competence in the information systems appeared since 2000, through the management planning of employment and competences. This is a process of human resources which appeared more recently in the HRIS[18].

The management of competences on the level of the HRIS is based on structures in the form of a thesaurus [19].

C.A.4 *Standardization HR-XML :*

Consortium HR-XML for human resources was created in 1999 on the initiative of a hundred companies resulting from 22 countries in the related services with the human resource management as well as companies of software solutions [20]. It is an organization which proposes a set of XML specifications to facilitate the automated exchange and the treatment of the numerical information related to the human resource management. The model of competence presented suffers from the problems involved in diversities of the terminologies used by the corporate.

II. RESULTS AND DISCUSSION

This review of the literature shows that the definition of competence and the characterization are not stabilized yet. Competence is a resource made up of elements which are non-elementary.

We retain the common elements which allow a characterization of the concept of competence identified by [2], [21]:

i. Competence is intrinsically related to an actor.
ii. Competence is expressed in a situation of a given work.
iii. Competence is a dynamic combination of cognitive resources and resources of work environment.
iv. Competence is regarded as a provision to act.

A typology of competences was proposed by [7] covering mainly the principal elements of characterization mentioned below:

v. Technical competence (Know-how)
vi. Behavioral competence (To know To be) are represented by the following components:
vii. Behavior of the collaborator as an individual and his relationship in work (Work, Personality)
viii. Behavior of the collaborator in relation to the others (Communication, Management)
ix. Behavior of the collaborator compared to the organization (Spirit of membership)

Another typology is proposed by [17]. This approach structures the competencies as follows:
Knowledge:
Theoretical knowledge
Current knowledge
Procedural knowledge

Know-how:
Formalization knowhow

Empirical knowledge
Behavioral competence (knowledge to be)
Relational knowledge
Cognitive knowledge
Behavior

Whatever the detail suggested of the components of competence, it cannot be generalized for a set of establishment. In this direction, we propose a generic model based essentially on the following concepts:

i. According to the definition of competence adopted by the company (knowledge, know-how, Attitude...), the competence is broken down into elements.
ii. These elements are then broken down into tangible and measurable components.
iii. The evaluation of the components is carried out according to the level of mastery of this component by the employee, this is the case of the acquired competence or the level of proficiency required to occupy a workstation this is the case of the required competence.
iv. The competences are structured in field and under field of competence.
v. Competence evolves over time.
vi. Job is a practice and knowledge recognized in the professional world, independent of the Organization.
vii. Employment is the exercise of a job in an organization.
viii. The workstation is the employment corresponds to the mobilization of the competences for the realization of the tasks.

Through this description of the competency management terms, we conclude that these terms are linked primarily by inheritance relationships. The ontology therefore lends itself well to conceptually structuring the competence.

Regardless of such treatment of competences, we identified few contributions based on formal approaches to characterize the competences held by the employees and the work stations. The relational data bases are frequently used as structure of data which can involve a loss of the semantics of the data. The thesauri used by the information systems of human resources present a small degree of formalization in the presentation of knowledge [22] (the field of knowledge is structured on the basis of hierarchical conceptual relations, equivalence and associative).

The development dedicated to the management of competences offer the tools which use references of a skill and of a job. However, each company is different and it is very difficult to be able to preserve a reference frame of a company to another. We can conclude that semantics Web technologies remain the most suitable solutions for the treatment of competences. The field of the semantic Web offers panoply of standards to represent knowledge and to handle them. Competence is a field of knowledge with excellency.

The whole of this knowledge will be represented in the form of ontology. This latter will be exploited with the external resource WordNet as supports in order to search of the having candidates competences necessary for a given job. The algorithm K-nearest neighbour (k-NN) is chosen. It is largely used for the technique search [23]. We will briefly present in the following the KNN classification method as well as the ontology.

A. KNN Classification

Classification is the process of finding a model (or function) that describes and distinguishes data classes or concepts [24]. The model is derived based on the analysis of a set of training data (i.e., data objects for which the class labels are known). The model is used to predict the class label of objects for which it is unknown. The derived model may be represented in various forms, such as classification rules, decision trees, mathematical formulae, or neural networks [24]. The Nearest-neighbour (KNN) is a advanced technique for data classification.

Nearest-neighbour classifiers are based on comparing a given test tuple with training tuples that are similar to it. The training tuples are described by n attributes and are stored in an n-dimensional pattern space. When given an unknown tuple, a k-nearest-neighbor classifier searches the k training tuples that are closest to the unknown tuple. These k training tuples found are the k "nearest neighbors" of the unknown tuple.

The Closeness is defined in terms of a distance metric, such as Euclidean distance. Usually, with continuous features, the Euclidean distance between two points or tuples tuples, say, X1 (x11, x12, …. , x1n) and X2 (x21, x22, …. , x2n), is [24]

$$dist(\boldsymbol{X}_1, \boldsymbol{X}_2) = \sqrt{\sum_{i=1}^{n} (x_{1i} - x_{2i})^2}.$$

If there is only one feature, the Euclidean distance is the absolute value of the difference.

$$D = \sum_{i=1}^{n} |x_i - y_i|$$

Manhattan distance is also used for continuous variables.

Typically, it is usually recommended to standardize the feature variables if their units are quite different. Likewise, it would be given more weight to feature variables with larger units. If k = 1, the object is simply assigned to the class of its very nearest neighbor.

For categorical features (gender, or age classified as young, middle-aged, and old) other metrics can be used. The overlap metric (also called the Hamming distance)[25]. The

distance between two values is 0 if they are the same and 1 if they are different.

KNN is a simple algorithm that stores all available cases and classifies new cases based on a similarity measure KNN is conceptually simple, yet able to solve complex problems.

B. ONTOLOGY

The word ontology is resulting from the philosophy field. In the last decade, ontologies are widely used in Knowledge Engineering, Artificial Intelligence and Computer Science. for this, we have read many definitions about what an ontology is and have also observed how such definitions have changed and evolved over the years. One of the first definitions was given by Neches and colleagues [26] "An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary". This definition, gives us vague guidelines to build an ontology and indicates that an ontology includes not only the terms that are explicitly defined in it, but also the knowledge that can be inferred from it. Some years later, Gruber defined ontology [26] "An ontology is an explicit specification of a conceptualization". This definition became the most quoted in scientific literature and by the ontology community. Many definitions based on Gruber's definition, have taken place since then. For exemple Borst [26] had modified slightly Gruber's definition as follows: "Ontologies are defined as a formal specification of a shared conceptualization". Wide variety of languages for explicit specification of ontology have been proposed by The World Wide Web Consortium (W3C), aim to represent the knowledge contained in an ontology in a simple and human-readable way and machine and to simplify the handling of resources. The best-known languages come from the W3C, such as RDF (Resource Description Framework), RDF S (Resource Description Framework Schema), OWL (Web Ontology Language). The RDF data model makes statements about resources expressions, known as triples, which follow a "subject–predicate–object" structure. The subject denotes the resource, and the predicate denotes traits or aspects of the resource, and expresses a relationship between the subject and the object. Also the SPARQL language is designed to query RDF which is described by a set of specifications from the W3C.

C. IMPLEMENTATION

The ontology is designed on the basis of the model proposed for the characterization of the competence through the typology of these components and their scale of evaluation was translated in an operational language of definition of ontologies. Our choice concerns the language OWL for the codification of the ontology, because it has a wealth in semantic features. We implement this ontology at the level of the editor "Protégé "(Fig.1).

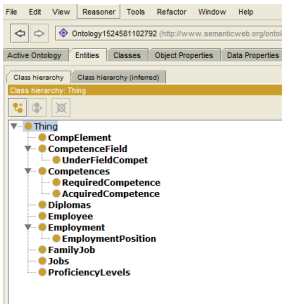


Figure: 1

This ontology was coupled with the lexical data base WordNet for searching candidate to a vacancy with the appeal of the algorithm K-Nearest Neighbors (KNN). WordNet is used of the synonymy treatment (synonyms are words that have similar meanings). We chose the programming language R because of its programming features. R is a free software environment for statistical computing, rich in Statistical functions which are indispensible for data mining. The R Core Team is now responsible for development and maintenance of R. R was ranked number 1 in the KDnuggets 2016 (a leading site on Business Analytics, Big Data, Data Mining, Data Science, and Machine Learning).
The plan of this application is raised in the Fig. 2.

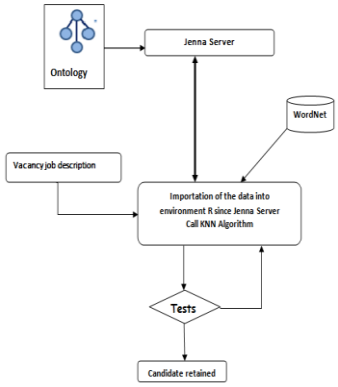


Figure: 2

The application of KNN (k – Nearest neighbor) algorithm using R programming uses the knn function of the FNN package. For this we proceed as follows:

i. Creating training and test data set.
ii. Defining an attribute to characterize the classes: each instance is assumed to belong to a class.
iii. Choice of a value of k
iv. Input : instance y
v. Determine the instances k neighbors to the instance Y.

D. TESTS

The matrix of instances retrieved from the ontology coupled with the Wordnet database for synonyms treatment of the attribute called "nom". This attribute identifies the classes. The attributes of matrix are essentially the skills

acquired (elements of skills and their levels: evol, adap, eff, ….) for an area of competence. For the case or k=3 and the training data are the same as for test, the result is as follows (Fig. 3)

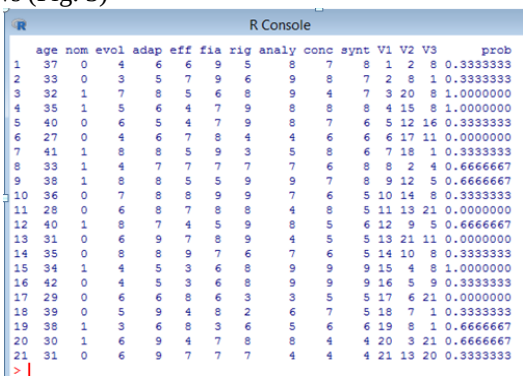


Figure: 3

The levels of control of competence are as follows: evol=8, adap=8, Effi=5 fiab=5, rig=8, ana=8, conc=7 and syn =6. We will search the employees having competences nearest with the necessary competences. The result got by replacing the data test by those of necessary competences arises as follows. (Fig. 4)

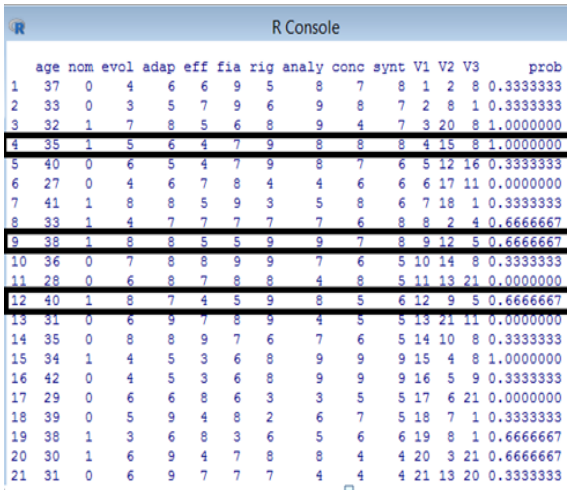


Figure: 4

III. CONCLUSION AND PERSPECTIVE

The recognition of the important role of the competence in the organizations management has made the administration of the competence a topic of interest to professionals as well as several research teams. It is not a simple storage and archiving of data but rather a structuring of the knowledge relating to the different situations of the employee's life in the company through models. Semantic Web technologies can make interesting contributions to component management as is the case that has been presented.

IV. REFERENCES

[1] Guide Méthodologique de support de la Gestion Prévisionnelle des Emplois et des Compétences Royaume du Maroc et Banque Mondiale 2008

[2] Belkadi F et al. "Contribution au pilotage des compétences dans les activités de conception : de la modélisation des situations à la caractérisation des compétences" 2006

[3] Reinbold et Breillot "Gérer la compétence dans l'entreprise" Préface de Anne Dousset Dynamiques d'entreprises ECONOMIE 2000

[4] Le Boterf Guy "Ingénierie et évaluation des compétences", et repenser la compétence éd. d'Organisation 2002

[5] Mc Clelland D. C., "Testing for Competence" 1973

[6] Ley T., "Organizational Competency Management: A Competence Management Approach", Shaker verglas 2006.

[7] Foveau "Référentiels des compétences et des métiers : une approche ontologique " Thése de Doctorat en informatique Université de Savoie et Université de Lausane 2007

[8] Bloom B.S., Krathwohl D.R., "Taxonomy of Educational Objectives: The classification of Educational goals, Handbook I: cognitive omain" Longmans New York, USA, 1956.

[9] Pépiot G., "Méthodologie des entreprises sur la base des compétences", thèse de doctorat en génie mécanique, Ecole Polytechnique Fédérale de Lausanne, EPFL 2005.

[10] Michel S., Ledru M., "Capital compétence dans l'entreprise : une approche cognitive", éd. ESF, 1991.

[11] Bourland K. E., Carl L. K., "Parallel-Machine sheduling with Fractional operator requirement", IIE Transactions, Vol. 26, N° 05, 1994

[12] De Korvin A., Shipley F., Kleyle R., "Utilizing fuzzy compatibility of skill sets for team selection in multi-phase projects", Journal of Engineering Technology Management, vol 19, 2002

[13] Canos L., Liern V., "Some fuzzy models for human resource management", Int. Journal of Technology, Policy and Management, Vol. 4, N° 4, 2004

[14] Imane Boughedda, Samia Djamai,Saloua Chettibi " Un Algorithme d'Appariement des Compétences pour 'Affectation des Modules d'un Parcours Informatique aux Enseignants " 2011

[15] Zouhel Boucetta, , Zizette Boufaida et Leila Yahiaoui "Appariement sémantique des documents à base d'ontologie pour le e-Recrutement "

[16] Michel Authier et Pierre Lévy "Les arbres de connaissances Par Collection " : Poche / Essais Éditeur: La Découverte 1999

[17] Harzallah, M., Berio, G. & Vernadat, F., "Modelling and Analysis of Individual competencies to improve industrial performances". IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans, 2006

[18] Bernard Just "Pas de DRH sans SIRH Présentation, métiers, technologies, méthodologie de mise en place " Editeur : Editions Liaisons collection : Entreprise et Carrières 2012

[19] Michelle G., Patrick G. "SIRH Système d'information des ressources humaines" Collection : Management Sup, Dunod Parution : 2010

[20] Jeans-Francois Borgne "La norme HR XML Mémoire présenté pour examen probatoire en Informatique" CNAM 2005

[21] Julien GIRODON "Proposition d'une approche d'amélioration des performances des organisations par le management opérationnel de leurs connaissances et compétences" Thése de doctorat L'Université de Loraine Spécialité : génie des systémes industriels 2015

[22] Thérèse Libourel. "Journées Ontologie en Sciences Humaines et Sociales" 9 - 10 novembre 2015 – Tours

[23] Khadim DRAME "Contribution à la construction d'ontologies et à la recherche d'information: application au domaine medical " thése de doctorat spécialité Informatique et Santé L'Université de Bordeaux décembre 2014.

[24] Jiawei Han University of Illinois at Urbana–Champaign Micheline Kamber Jian Pei Simon Fraser University Data Mining Concepts and Techniques Third Edition Morgan Kaufmann Publishers 2012 by Elsevier

[25] Johannes Ledolter Department of Management Sciences Tippie College of Business University of Iowa DATA MINING AND BUSINESS ANALYTICS WITH R 2013 John Wiley

[26] Asunción Gómez-Pérez, Mariano Fernández-López and Oscar Corcho Ontological Engineering with examples from the areas of Knowledge

Management, e-Commerce and the Semantic Web Springer 2004

[27] doctorat spécialité Informatique et Santé L'Université de Bordeaux décembre 2014.