

Appariement sémantique des documents à base d'ontologie pour le e-Recrutement

Zouhel Boucetta¹, Zizette Boufaïda¹ et Leïla Yahiaoui^{1,2}

¹ Laboratoire LIRE – Université Mentouri Constantine- 25000, Algérie

[boucetta.zouhel, zboufaïda, leila.yahiaoui}@gmail.com](mailto:{boucetta.zouhel, zboufaïda, leila.yahiaoui}@gmail.com)

² Laboratoire LIRIS – Université Claude Bernard Lyon 1

F-69622 Villeurbanne cedex, France

leila.yahiaoui@liris.cnrs.fr

Résumé : le Web a permis une évolution importante du marché du recrutement en ligne. Cependant, sur le web actuel, les outils de modélisation et de recherche des CVs/offres d'emploi sont souvent syntaxiques et limités. Afin d'automatiser et d'optimiser le processus de recrutement dans le cadre du web, nous proposons dans cet article, une idée d'appariement sémantique entre les requis d'un recruteur et les acquis d'un chercheur d'emploi, exprimés sous forme d'annotations sémantiques. L'annotation est utilisée afin de permettre à un utilisateur (chercheur d'emploi/recruteur) de décrire le contenu sémantique de son document (CV/offre d'emploi) à l'aide du vocabulaire d'un référentiel commun (ontologie), inspiré des parties communes entre ces documents (exigence, qualification personnelle, diplôme, métier et compétences sous-jacentes). La compétence constitue l'élément crucial de cette modélisation et permet ainsi un appariement plus pertinent.

Mots-clés : Web sémantique, e-recrutement, ontologie, compétence, annotation sémantique, appariement sémantique, CV/offre d'emploi.

1 Introduction

L'Internet s'est transformé rapidement en un outil de recrutement performant. 53% des recrutements en Allemagne sont le résultat d'une postulation en ligne [3] ; un chiffre qui est susceptible d'augmenter. Cependant, vu les faiblesses des outils dédiés à la gestion et à la recherche des CVs/offre d'emploi, l'écoulement de l'information sur ce marché de travail est loin d'être optimal. Une solution peut être trouvée dans le Web sémantique qui vise à enrichir les structures syntaxiques du Web actuel avec leur contenu sémantique. Par conséquent, l'exploitation de ses technologies dans le cadre du recrutement électronique semble être bénéfique, surtout pour assurer un rapprochement automatique entre les offres et les demandes d'emploi. Le principe consiste à expliciter le contenu de ces documents à l'aide d'éléments ontologiques formant un référentiel commun (ontologie), inspiré des parties communes entre ces documents (diplôme, expérience professionnelle, compétence, qualification personnelle, exigence). Ce référentiel peut être utilisé par un recruteur (ou un chercheur d'emploi) pour annoter son document avec ses requis (ses acquis). Les méta-données résultantes peuvent être exploitées pour automatiser le processus de rapprochement entre son document et les CVs (offres d'emploi) disponibles. Dans cet

article, nous proposons des techniques d'appariement sémantique entre CVs et offres d'emploi, dans lesquelles l'appariement des compétences est prépondérant. L'article est organisé comme suit : la section 2 présente un état de l'art concernant les travaux existants dans ce domaine. La section 3 présente l'architecture globale du système proposé dans laquelle nous retrouvons essentiellement un composant ontologique, une interface d'annotation et un composant d'appariement. La section 4 décrit le modèle de la compétence proposée. La section 5 présente l'ontologie développée pour modéliser le contenu sémantique des documents, et qui sera utilisée dans les processus d'annotation et d'appariement. Dans la section 6, nous décrivons les algorithmes d'appariement sémantique à appliquer entre les documents et qui traitent trois aspects : i) les qualifications personnelles, ii) les exigences, iii) les compétences. Nous terminons l'article par une conclusion et des perspectives.

2 Etat de l'art

Plusieurs travaux concernant le e-recrutement se basent sur la technique d'annotation sémantique, du fait de sa simplicité et de sa capacité à expliciter le contenu des documents. L'apport réside dans le fait que les données formalisées produites seront utilisées dans un rapprochement automatique entre ces documents.

Dans ce cadre, le projet BRIDGE [8] vise à traiter les problèmes du e-recrutement selon une approche basée sur la gestion des compétences. L'idée consiste à permettre à un demandeur d'emploi (ou un recruteur) d'identifier et de représenter formellement les compétences sous-jacentes à son CV (ou son offre d'emploi). Ces compétences sont utilisées pour raffiner le processus d'appariement entre l'offre et la demande d'emploi. Le modèle de compétence proposé est bien détaillé, mais les ontologies utilisées et les algorithmes d'appariement à appliquer sont en cours de développement.

Bizer et al [2] à leur tour, proposent le développement d'une ontologie de gestion des ressources humaines, par l'intégration de normes et de classifications existantes. Leur algorithme d'appariement basé compétence semble être pertinent, mais l'ontologie est générale et le modèle de compétence sous-jacent à l'ontologie n'a pas été défini.

Dans une première version de notre travail, Yahiaoui et al [9], [10] ont développé une ontologie détaillée pour la modélisation du contenu des CVs et des offres d'emploi, explicitant un modèle de compétence relativement simple. Cette simplicité visait à simplifier le processus de rapprochement à appliquer ultérieurement entre ces documents. Néanmoins, la négligence du contexte dans le modèle de la compétence a engendré quelques faiblesses dans le processus d'appariement. Pour mieux rapprocher ces documents d'un point de vue de la compétence, nous introduisons dans la modélisation de cette dernière, des notions concernant le contexte organisationnel et économique dans lequel cette compétence est acquise ou requise, afin que les résultats du processus d'appariement basé compétence soient les plus faibles possibles (à titre d'exemple, deux demandeurs d'emploi ayant la même compétence et le même niveau dans cette compétence, seront distingués par le contexte s'ils ont travaillé dans des entreprises de taille différentes « contexte économique »). De plus, nous considérons dans ce travail, des poids de pondération à appliquer aux exigences (type de contrat, salaire, etc.), aux compétences (sous-jacentes aux métiers/diplômes) et aux qualifications personnelles (âge, sexe, etc.) annotant le document de l'utilisateur (recruteur ou chercheur d'emploi), et qui seront explicitement spécifiés par ce dernier afin d'exprimer son choix.

3 Le système proposé

L'objectif de notre travail consiste à proposer un système supportant l'appariement sémantique entre les offres d'emploi et les Cvs dont l'architecture est présentée par la fig 1. Dans ce système : (1) un utilisateur peut déposer son document (CV/offre) sur un serveur de documents ; (2) il pourra utiliser l'interface d'annotation pour exploiter l'ontologie-ER (3) et générer des méta-données décrivant le contenu de son document (4). Un utilisateur peut aussi envoyer une requête à l'interface d'appariement (5) afin de calculer le degré de rapprochement de son document avec ceux disponibles (de type opposé) sur le serveur. Cette interface formalise la requête de l'utilisateur et l'envoie au composant d'appariement (6) utilisant un raisonneur (7), exploitant l'ontologie-ER (8) et l'entrepôt des annotations liées aux documents (9). Le résultat retourné (10) est un ensemble de quadruplets (URI, $DA_{QPersonnel}$, $DA_{Exigence}$, $DA_{Compétence}$) tels quels que : URI est l'identifiant de l'un des documents rapprochés, $DA_{QPersonnel}$ le degré d'appariement de son document avec ce dernier selon les qualifications personnelles, $DA_{Exigence}$ le degré d'appariement selon les exigences et $DA_{Compétence}$ le degré d'appariement selon les compétences sous-jacentes. Ce résultat est retourné par la suite à l'utilisateur (11) afin qu'il puisse décider.

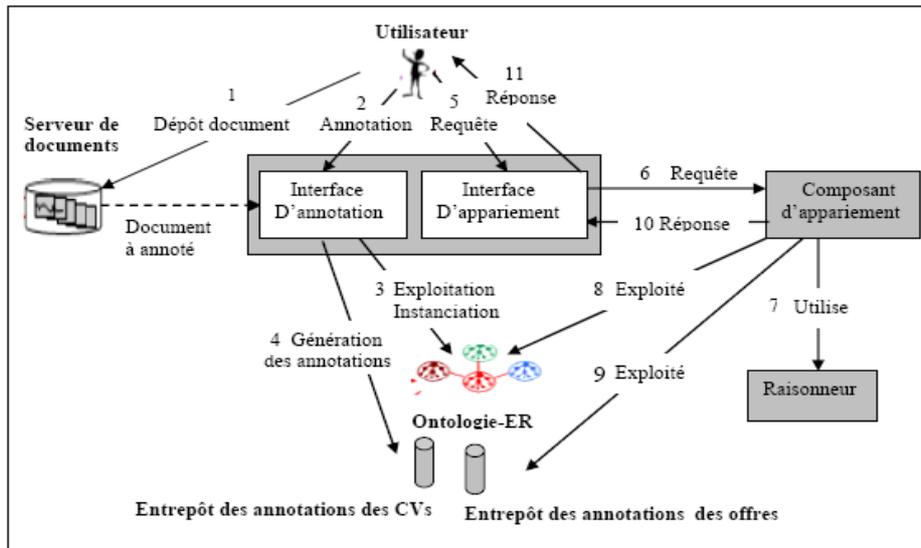


Fig. 1. Architecture du système

4 Modélisation de la compétence

Une compétence est identifiée comme un ensemble de connaissances mises en action dans la réalisation des tâches quotidiennes de l'entreprise. Elle se manifeste sous forme d'un comportement et peut être scientifique et technique ou une aptitude comportementale [9]. Le modèle de la compétence que nous proposons est un enrichissement du modèle proposé dans [9] par le concept de contexte [8] et la notion de ressources (savoir, savoir-faire, habilité) [7]. La figure 2 présente ce modèle, basé

sur la définition suivante : une compétence peut être une *aptitude comportementale* ou une *compétence scientifique ou technique*. Les aptitudes sont classées selon des catégories [3], alors que la compétence scientifique/technique est caractérisée par la mobilisation d'un ensemble de *ressources* pouvant être un *savoir*, un *savoir-faire* ou une *habileté* (un don). Chaque savoir ou savoir-faire est lié à un *thème* particulier, appartenant à une hiérarchie de thèmes *généraux* ou *spécifiques* à un domaine (le domaine informatique et télécommunication dans notre cas).

La compétence est également caractérisée par un *niveau d'expérience*, dont la valeur dépend du type de celle-ci. Dans le cas d'une compétence scientifique ou technique, ce niveau peut avoir l'une des valeurs suivantes : Débutant (D/1), Intermédiaire (I/2), Moyen (M/3), Compétent (C/4) ou Professionnel (P/5) ; alors qu'il peut avoir juste l'une des deux valeurs, forte (FO/2) ou faible (FA/1), dans le cas d'une aptitude. Le contexte de la compétence sous-entend le *contexte organisationnel* et *économique*. Le contexte organisationnel représente une évaluation du groupe de formation (cas d'une compétence mobilisée par un diplôme) ou du service de travail (cas d'une compétence mobilisée par un métier), selon le nombre des personnes engagées, il peut être {grand service (6), service moyen (5), petit service (4), petit groupe (3), groupe moyen (2), grand groupe (1)}. Le contexte économique décrit plutôt l'importance de l'organisme formateur (cas d'un diplôme) ou de l'organisme recruteur (cas d'un métier), selon la réputation ou l'importance économique de cet organisme (chiffre d'affaire de la société, qualification d'une université parmi d'autres, etc.). Un contexte économique peut être {Grande entreprise (6), entreprise moyenne (5), petite entreprise (4), grande école (3), université (2), centre universitaire (1)}.

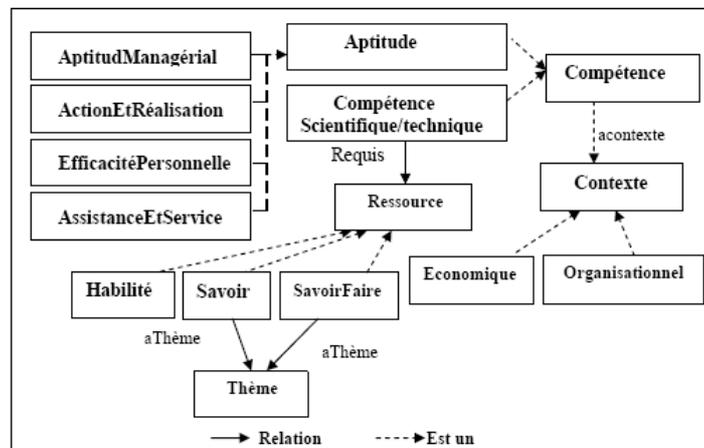


Fig. 2. Le modèle de la compétence proposée

5 Une ontologie modélisant le contenu sémantique

Les ontologies sont cruciales dans le cadre du e-recrutement car elles permettent aux recruteurs et aux chercheurs d'emploi de partager un référentiel commun pour décrire le contenu de leurs documents d'une façon précise, sémantique et formelle. Ce

référentiel servira comme support pour le processus d'annotation générant des données formalisées, permettant ainsi un traitement automatique dans le cadre d'un rapprochement entre les CVs et les offres d'emploi. Dans cette perspective, nous avons construit une ontologie nommée « ontologie-ER » dont les concepts et les relations sont inspirés des parties communes les plus significatives pour les CVs et les offres d'emploi (qualification personnelle, exigence, diplôme, expérience professionnelle, compétence).

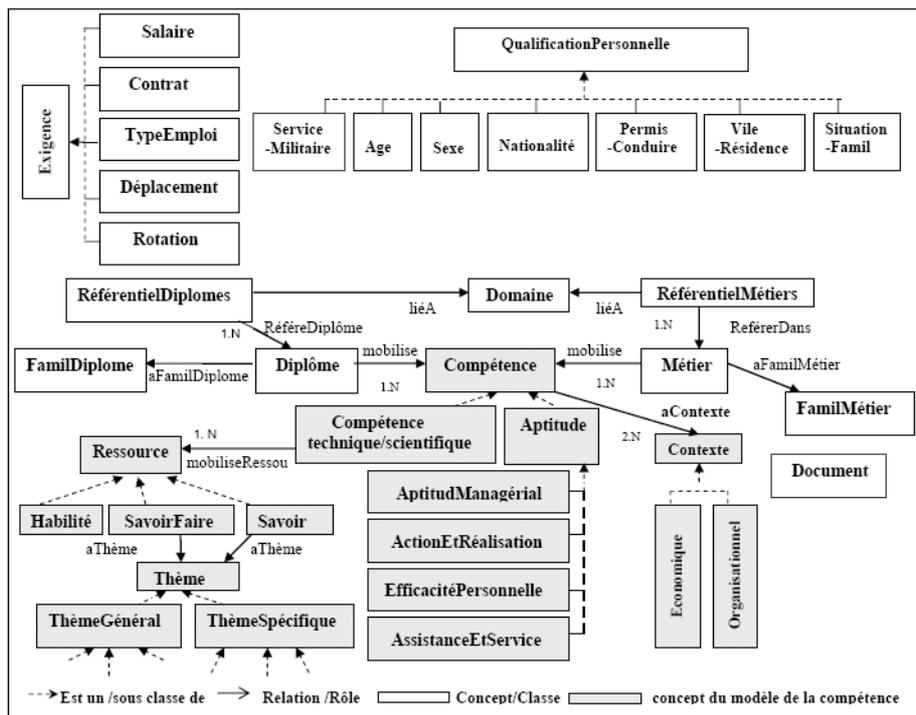


Fig. 3. Architecture détaillée de l'ontologie-ER

5.1 Description de l'ontologie-ER

La figure 3 présente les éléments de cette ontologie, que nous décrivons brièvement :

- *Document* : c'est le concept décrivant le document à annoter (CV ou offre) à travers son nom, son type, son URI, etc.
- *Qualification personnelle* : c'est un concept décrivant les caractéristiques personnelles les plus importantes, qu'un recruteur peut exiger ou qu'un candidat peut avoir. Il s'agit du sexe, de l'âge, de la nationalité, de la ville de résidence, etc.
- *Exigence* : ce concept décrit certaines contraintes imposées par le recruteur ou le chercheur d'emploi telles que : le salaire, le type de contrat, le déplacement, etc.
- *Diplôme* : c'est un concept qui décrit les diplômes appartenant à une famille particulière « FamilDiplôme ». Un référentiel pour ces diplômes

« RéférentielDiplomes », concernant un domaine particulier « Domaine », a été inspiré de la réforme LMD [4] de l'enseignement supérieur algérien. Le concept diplôme est lié également au concept « Compétence » pour pouvoir attester les compétences mobilisées par un diplôme particulier.

- *Métier* : c'est un concept qui décrit les métiers appartenant à une famille de métiers « FamilMétier ». Un référentiel pour ces métiers « RéférentielMétiers », concernant un domaine particulier « Domaine », a été inspiré de CIGREF [3] et de O*net [6]. Le concept métier est lié également au concept « Compétence » pour pouvoir attester les compétences mobilisées par un métier particulier.
- *Compétence* : c'est un concept décrivant la compétence selon le modèle proposé dans la section 4. Ses instances sont liées généralement à un métier ou à un diplôme particulier.

5.2 Annotation des documents à base de l'ontologie-ER

L'annotation sémantique est l'assignement des éléments d'une ontologie (instance de classe, instance de relation) aux éléments d'un document. Un recruteur ou un chercheur d'emploi peut exploiter l'ontologie-ER pour annoter sémantiquement son document CV/offre, à travers un outil d'annotation existant.

Certaines rubriques dans l'interface d'annotation ont des valeurs prédéfinies (métiers, diplômes, compétences sous-jacentes à ces métiers et diplômes).

D'autres rubriques doivent être renseignées par l'utilisateur. Ce dernier décrit les qualifications personnelles acquises (cas du CV) ou requises (cas de l'offre). Pour chaque requis ou acquis, un poids d'importance doit être fixé par l'utilisateur ainsi que d'autres informations importantes telles que la distance tolérée par le recruteur (ou chercheur d'emploi) concernant la ville de résidence du chercheur d'emploi. Les exigences sont définies en termes de contraintes imposées par le recruteur (ou le chercheur d'emploi) (telles que le montant du salaire, le déplacement, etc.), L'utilisateur assigne également un poids à chacune de ces contraintes selon son importance pour lui. D'autres données telles que la différence de salaire tolérée par un recruteur (ou chercheur d'emploi) entre sa proposition et celle exigée par un chercheur d'emploi (ou recruteur) est également annotée.

Finalement, l'utilisateur sélectionne explicitement des compétences (ainsi que les niveaux et les contextes de celles-ci) sous-jacentes à un métier ou un diplôme figurant dans son document, via un choix dans une liste prédéfinie, inspirées des différents référentiels de métiers et de formations utilisés. Il doit également assigner un poids à chacune de ces compétences selon son importance pour lui.

Le résultat de ce processus est un ensemble d'annotations réparties sur deux entrepôts d'annotations (l'entrepôt des CVs et celui des offres d'emploi).

6 Appariement sémantique entre CVs et offres d'emploi

L'appariement sémantique des requis d'un recruteur (fichier d'annotations d'une offre d'emploi : FA_{offre}) et les acquis d'un chercheur d'emploi (fichier d'annotations d'un CV : FA_{CV}) (ou vice versa) revient à calculer leur degré de rapprochement selon les trois critères : qualifications personnelles, exigences et compétences. Nous proposons trois algorithmes d'appariement. Pour les compétences, nous avons proposé deux techniques d'appariement, exact ou rapprochée.

6.1 Appariement des qualifications personnelles

Les qualifications personnelles constituent la première catégorie d'informations à traiter. Pour illustrer leur appariement, nous considérons des éléments tels que l'âge, la ville de résidence, la nationalité, le sexe, etc. Chaque qualification personnelle requise (R) annotée dans l'offre (FA_{offre}) est recherchée avec exactitude dans l'ensemble des acquis annotés dans le CV (FA_{cv}). Dans le cas d'une égalité avec un acquis de même type, le poids d'importance du requis est cumulé. Pour le requis « ville de résidence », le recruteur limite la distance acceptée entre la ville requise et celle de résidence d'un chercheur d'emploi par une distance maximale acceptée. Si cette différence est inférieure à la distance maximum tolérée, elle servira dans l'ajustement du poids du requis avant que ce dernier ne soit cumulé. Le degré d'appariement des qualifications personnelles est le rapport entre le cumul des poids des requis vérifiés et le cumul des poids de toutes les qualifications requises par le recruteur. La fonction (Distance) est utilisée dans l'algorithme de l'appariement des qualifications personnelles pour calculer la distance entre deux villes. L'algorithme de l'appariement des qualifications personnelles est le suivant :

```

AQPersonnel ( $FA_{\text{offre}}, FA_{\text{cv}}$ )
  cumul : le cumul des poids des requis vérifiés par le chercheur d'emploi ;
  QPersonCV : les qualifications personnelles extraites du CV ;
  QPersonOffre : les qualifications personnelles extraites de l'offre ;
  cumul_tot : le cumul des poids des requis ;
  DAQPersonnel : le degré d'appariement des qualifications personnelles ;
  QPersonCV := ∅ ; QPersonOffre := ∅ ; cumul := 0 ; cumul_tot := 0 ; DAQPersonnel := 0 ;
  Extraire ( $FA_{\text{cv}}, FA_{\text{offre}}, QPerson_{\text{CV}}, QPerson_{\text{offre}}$ ) ;
  Pour chaque requis  $R \in QPerson_{\text{Offre}}$  Faire
    cumule_tot := cumul_tot + poids (R) ; /*un acquis est symbolisé par A*/
    Si  $\exists$  acquis  $A \in QPerson_{\text{CV}}$  / type (A) = type (R)  $\wedge$  valeur (A) = valeur (R) alors
      cumul := cumul + poids (R) ;
    Sinon /* DAcc est la distance maximale acceptée entre la ville requise et acquise */
      Si (type(R) = type (A) = 'VilleR')  $\wedge$  (distance(R, A) < MaxDAcc) alors
        cumul := cumul + (poids(R)) * (1 / MaxDAcc) * (distance(R, A)) ;
      Fin si
    Fin si
  Fin pour
  DAQPersonnel := (cumul / cumul_tot) * 100 ;
Fin

```

6.2 Appariement des exigences

Les exigences constituent la deuxième catégorie d'informations à traiter. Pour illustrer leur appariement, nous considérons des éléments tels que la rotation (3x8), le type d'emploi, le salaire, etc. Chaque exigence du recruteur (R) annotée dans le fichier des annotations de l'offre (FA_{offre}), est recherchée avec exactitude dans l'ensemble des exigences d'un chercheur d'emploi annoté dans le fichier des annotations du CV (FA_{cv}). En cas d'égalité avec une exigence de même type, le poids d'importance du requis est cumulé. Le recruteur peut aussi limiter la différence acceptée entre « le salaire » qu'il exige et « le salaire » qui peut être demandé par un chercheur d'emploi par une différence maximale acceptée. Cette

dernière sera utilisée pour ajuster le poids à cumuler. Le degré d'appariement des exigences est le rapport entre le cumul des poids des requis vérifiés et le cumul des poids de toutes les exigences annotées par le recruteur. L'algorithme de l'appariement des exigences suit le même principe que l'algorithme d'appariement des qualifications personnelles.

6.3 Appariement des compétences

Vu l'importance des compétences dans la sélection du chercheur d'emploi le plus qualifié ou de l'offre d'emploi la plus intéressante, nous avons proposé deux techniques d'appariement des compétences : l'appariement exact (CompExact) et l'appariement rapproché (RapprochéComp).

6.3.1 Principe de l'appariement exact des compétences

L'appariement exact est appliqué dans le cas où le recruteur cherche un candidat qui possède exactement les compétences requises pour un poste. Une compétence scientifique ou technique requise (R) est recherchée avec exactitude dans l'ensemble des compétences acquises (Comp_{CV}). Le cas échéant, nous nous intéressons aux compétences acquises ayant un niveau supérieur ou égal à celui de la compétence requise et un contexte meilleur sinon similaire à celui du requis. Le même principe est appliqué aux aptitudes.

La fonction (Eva) est utilisée dans l'algorithme de l'appariement exact des compétences. Elle retourne le cumul de la valeur des contextes organisationnel et économique. A titre d'exemple, pour une compétence acquise dans un grand service d'une grande entreprise, Eva (Context (compétence_acquise)) vaut la valeur du contexte organisationnel d'une grande entreprise + la valeur du contexte économique d'un grand service = 6+6=12.

```

CompExact (FAOffre, FACV)
  cumul : le cumul des poids des requis vérifiés par le chercheur d'emploi ;
  CompCV : les compétences extraites du CV;
  CompOffre : les compétences extraites de l'offre d'emploi ;
  cumul_tot : le cumul des poids des requis ;
  DACompétence : degré d'appariement des compétences ;
  CompCV := ∅; CompOffre := ∅; cumul := 0 ; cumul_tot := 0 ; DACompétence := 0 ;
  Extraire (FAOffre, FACV, CompOffre, CompCV) ; /*extraire les compétences*/
  Pour chaque compétence R ∈ CompOffre Faire /* un acquis est symbolisé par A */
    cumul_tot := cumul_tot + poids (R) ;
    Si ∃ A ∈ CompCV / ((type (A) = type (R)) ∧ (Désignation (A) = Désignation (R)) ∨
      (Niveau (R) <= Niveau (A)) ∧ (Eva (Context(R)) <= Eva (Context (A)))) alors
      cumul := cumul + poids (R) ;
    Fin Si
  Fin pour
  DACompétence := (cumul / cumul_tot) * 100 ;
Fin

```

6.3.2 Principe de l'appariement rapproché des compétences

Dans ce cas, le rapprochement n'est pas sévère. Chaque compétence requise de l'offre est recherchée dans l'ensemble des compétences acquises dans le CV (FA_{CV}), selon

ses ressources ou sa catégorie, son niveau et son contexte. Si la compétence requise est de type aptitude sa catégorie est recherchée avec exactitude dans l'ensemble des catégories des aptitudes annotant le cv (FA_{CV}). L'appariement d'une compétence scientifique ou technique requise R avec une compétence acquise A, revient à l'appariement des deux ensembles de ressources mobilisées par ces deux compétences, selon la formule (2) sachant que les ressources sont de trois types (habilité, savoir, avoir-faire). L'habilité (H) de la compétence requise (R) est recherchée avec exactitude dans l'ensemble des habilités de la compétence acquise (A). Si elle est vérifiée, un poids de 1 (tout ou rien est cumulé dans $DA_{habilité}$. Le rapprochement des savoirs (ou savoir-faire) d'une compétence requise (R) avec ceux d'une compétence acquise A revient au rapprochement de leurs thèmes.

Chaque thème d'un savoir (ou d'un savoir-faire) d'une compétence (R) est recherché dans l'ensemble des thèmes des savoirs (ou savoir-faire) d'une compétence acquise (A) de même type (général ou spécifique), en utilisant la formule de similarité des concepts (3). Le meilleur degré de similarité sera cumulé dans DA_{savoir} (ou dans $DA_{SavoirF}$). Le degré de rapprochement global entre la compétence requise (R) et la compétence acquise (A) est le rapport entre la somme de ($DA_{habilité}$, DA_{savoir} , $DA_{SavoirF}$) et le nombre de ressources qui constituent R.

La fonction SimN calcule la similarité de leurs niveaux par l'application de la formule (4), alors que la fonction SimC calcule la similarité de leurs contextes qui est le produit de la fonction SimCO pour le contexte organisationnel par la fonction SimCEC pour le contexte économique, toujours selon la formule (4).

Les formules suivantes sont utilisées dans l'algorithme appariement rapproché :

Soit :

- $c_1, c_2 \in C$, C est l'ensemble de toutes les instances de type « Compétence scientifique/technique »
- P : la relation « mobilise » qui relie une compétence c_i à un ensemble de ressources Re.
- $A_S(P, c_i)$ [5] : la fonction qui retourne l'ensemble des instances de type « Ressource » mobilisées par la compétence c_i , tel que :

$$A_S(P, c_i) = \{r_j : r_j \in Re \wedge P(c_i, r_j)\}. \quad (1)$$

$F(c_1, c_2, P)$ [5] est une fonction qui calcule la similarité de deux compétences en se basant sur la relation P « mobilise » qui les relie aux instances d'un autre concept « Ressource ». Soit $r_j \in A_S(P, c_1)$, et $b \in A_S(P, c_2)$ et $N = |A_S(P, c_1)|$: le nombre d'instance de cet ensemble.

$$F(c_1, c_2, P) = \frac{\sum_{j=1}^N \max \{sim(r_j, b) / b \in A_S(P, c_2)\}}{N} \quad (2)$$

La fonction *sim* concerne dans notre cas le calcul de la similarité entre deux ressources (r_j, b) de type « Ressource », pouvant être de type spécifique : « Savoir », « Savoir-faire » ou « Habilité ».

Si le type spécifique de r_j est différent de celui de b alors $sim(r_j, b)=0$ (on ne peut pas comparer un savoir à une habilité par exemple)

Si r_j et b ont le même type spécifique « Habilité » alors $\text{sim}(r_j, b)=1$ Sinon (r_j et b sont tous les deux des savoirs ou des savoirs-faire)
 $t_1 := r_j.a\text{Thème}$ ayant le type T_1 et $T_2 := b.a\text{Thème}$ ayant le type T_2 (T_1 et T_2 sont des sous-classes de « Thème ») si le type spécifique de t_1 est différent de celui de t_2 («ThèmeGeneral » ou « ThèmeSpécifique ») alors $\text{sim}(r_j, b) = 0$
 Sinon appliquer la formule 3 [1] :

$$\text{Sim}(r_j,b) = (2*\log \text{PROF}(\text{MRCA}(T_1,T_2))) / (\log \text{PROF}(T_1) + \log \text{PROF}(T_2)) \quad (3)$$

Sachant que :

- $\text{PROF}(T_1)$ est la profondeur du thème T_1 dans la hiérarchie des thèmes de « ThèmeGeneral » ou « ThèmeSpécifique ». $\text{PROF}(T_1) = (\text{nombre des sous classes de } T_1+1)/\text{nombre de classes de la hiérarchie du concept « ThèmeGeneral » ou « ThèmeSpécifique »}$.
- $\text{MRCA}(T_1, T_2)$ est l'ancêtre commun le plus proche entre T_1 et T_2 dans la hiérarchie de thèmes.

Pour le calcul de la similarité (SimN) de deux niveaux de compétence n_1 et n_2 nous utilisons la formule (4) [2] : α est le nombre de niveau, il est égal à 5 pour les compétences scientifique/technique et à 2 pour les aptitudes

Pour le calcul de la similarité (SimCEc) de deux contextes économiques ctx_1 et ctx_2 , on utilise la même formule avec $\alpha=6$ pour le nombre de contexte économique. Même chose pour le contexte organisationnel (SimCOR).

$$= \begin{cases} 1 - 1/\alpha (a_1 - a_2) & \text{Si } a_1 - a_2 \geq 0 \\ 1 & \text{Si } a_1 - a_2 < 0 \end{cases} \quad (4)$$

Afin de mieux comprendre le principe de l'algorithme d'appariement rapproché des compétences données plus loin, nous traitons une offre d'emploi annotée avec une compétence technique {compétence = java, niveau = expert, contexte organisationnel = grand service, contexte économique = grande entreprise, poids =1 (poids du requis java)} et un CV est annoté avec {compétence = pascal, niveau = moyen, contexte organisationnel = service moyen, contexte économique = grande entreprise}.

Chaque compétence mobilise un ensemble de ressources (un extrait de la hiérarchie des thèmes des ressources est illustré dans (fig.4)). Pour java (savoir mathématique, savoir-faire java, habilité de raisonner), pour pascal (savoir mathématique, savoir-faire pascal, habilité de raisonner). Supposons que ce sont les seules compétences qui existent pour le CV et pour l'offre d'emploi. Le calcul de leur similarité selon leurs contextes, leurs niveaux et leurs ressources est le suivant :

cumul := cumul+ poids (R) * Max(S_i) (cumul = 0, poids (R) = 1)

S₁ :=F(R, A, mobilise)*SimN (Niveau(R), Niveau (A))* SimC (Contexte(R), contexte (A)).

Nous appliquons la formule (2) pour le calcul de F(R, A, mobilise)

DA_{habilité} := Sim (raisonner, raisonner) = 1

DA_{savoir} := Sim (Mathématique, Mathématique) =1

DA_{savoirF} := Sim (savoir-faire pascal, savoir-faire java) = (2* log (6/100))/(log (1/100) +log (1/100)) = 0.61 (le calcul de similarité est effectué par l'application de la formule (4) qui calcul la similarité de deux concepts dans une hiérarchie où U, le nombre de concepts dans la hiérarchie des thèmes spécifiques vaut 100)

F(R, A, mobilise) := DA_{habilité} + DA_{savoir} + DA_{savoirF} / N = (1+1+ 0.61)/3 = 0.87 (nombre de ressources de java vaut 3).

Nous appliquons la formule (4) pour le calcul de SimN
SimN (niveau (java), niveau (pascal)) = Sim (expert, moyen) = Sim (5 ,3) = 0.66
 Nous appliquons la formule (4) pour le calcul de SimCOR et de SimCEc.
SimC (contexte (java), contexte (pascal)) = similarité contexte organisationnel *
 similarité contexte économique = **SimCOR** (grand service, petit service)* **SimCEc**
 (grande entreprise, grande entreprise) = SimCOR (6, 4)* SimCEc (6, 6) = 0.66*1 = 0.66.
cumul := 1* 0.87 *0.66 * 0.66 = 0.378 ; **cumul_tot** = 1
DA_{Compétence} := (0.378/1) *100 = 37.8 % (degré d'appariement)

```

CompRapproché(FAOffre, FACV)
  cumul : le cumul des poids des requis vérifiés par le chercheur d'emploi ;
  CompCV : les compétences extraites du CV;
  CompOffre : les compétences extraites de l'offre d'emploi ;
  SI : un tableau ;
  cumul_tot : le cumul des poids des requis ;
  DACompétence : degré d'appariement des compétences ;
  CompCV := ∅ ; CompOffre := ∅ ; cumul := 0 ; cumul_tot := 0 ; DACompétence := 0 ;
  Extraire (FAOffre, FACV, CompOffre, CompCV) ;
  Pour chaque compétence requise R ∈ CompOffre Faire /*un acquis est symbolisé par A*/
    cumul_tot := cumul_tot+ poids (R) ; I := 0 ; A=0 ; SI := 0 ;
    Si R est une compétence scientifique/technique alors
      Pour chaque compétence scientifique/technique A ∈ CompCV Faire
        /* F est calculé par la formule (2) et SimN par la formule (4)*/
        SimC( Context(R),context(A)) = SimCOR* SimCEc
        /* SimCOR , SimCEc sont calculées par la formule (4)*/
        SI:=F(R, A, mobilise)*SimN(Niveau(R),Niveau(A))*SimC(Context(R),context(A));
        I := I +1 ;
      Fin pour
      cumul := cumul+ poids (R) * Max(SI) ;
    Sinon /* les même formule (4) sont utilisée pour le calcul de SimCOR, SimCEc, SimN */
      Pour chaque aptitude A ∈ CompCV/ catégorié (R)= catégorié (A) Faire
        SimC(Context(R),context(A)) = SimCOR* SimCEc ;
        SI:= SimN(Niveau(R),Niveau(A))*SimC(Context(R),context(A));
        I := I +1 ;
      Fin pour
      cumul := cumul+ poids (R) * Max(SI) ;
    Fin Si
  Fin pour
  DACompétence := (cumul /cumul_tot) *100 ;
Fin
  
```

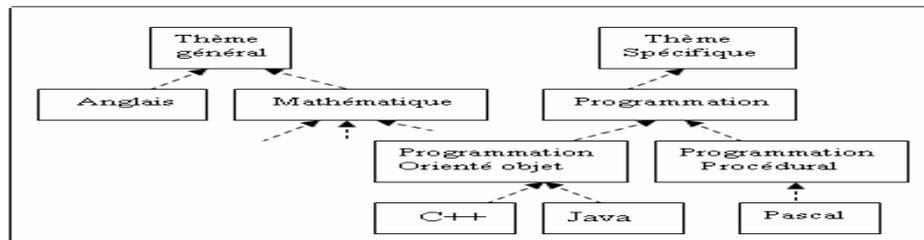


Fig.4. Extrait de la hiérarchie des thèmes des savoirs et des savoir-faire.

7 Conclusion

Dans ce travail, nous avons proposé une modélisation pertinente du contenu sémantique à base d'ontologie des CVs et des offres d'emploi, axée sur la modélisation de la compétence. Le modèle de la compétence adopté est riche parce qu'il tient compte du type de la compétence (scientifique et technique ou aptitude), des ressources relatives aux compétences scientifiques (habileté, savoir, savoir-faire), de la catégorie relative aux aptitudes (managériales, résolution des problèmes, ...), du niveau d'expérience et des contextes organisationnel et économique. Un outil d'annotation est utilisé pour enrichir les documents avec leur contenu. Des algorithmes d'appariement originaux permettent aussi bien le rapprochement basé sur les qualifications personnelles et les exigences, que celui basé sur les compétences sous-jacentes.

Pour augmenter la pertinence des résultats d'appariement, nous comptons prendre en compte des contextes plus complexes tels que le contexte technologique (la technologie d'un outil utilisé). Il serait également intéressant de considérer, dans la modélisation de l'ontologie, l'opérateur de formation, comme troisième acteur de notre système (portails de recrutement sémantique), en plus du recruteur et du chercheur d'emploi.

Références

1. Bernstein, A., Kaufmann, E., Buerki, C., Klein, M.: Object Similarity In Ontologies a Foundation for Business Intelligence Systems and High-Performance Retrieval. In: Twenty-Fifth International Conferences on Information Systems (2004)
2. Bizer, C., Heese, R., Mocho, M., Oldakowski, R., Tolksdorf, R., Eckstein, R.: The Impact of Semantic Web Technologies on Job Recruitment Processes. In: International Conference workshop on computer science (WI'05), (2005), <http://www.wiwiss.fu-berlin.de/suhl/radek/semmf>
3. Nomenclature, les emplois-métiers du système d'information dans les grandes entreprises (CIGREF), (2005), <http://www.cigref.fr/cigref/livelink.exe/Nomenclature>
4. Réforme LMD de l'enseignement Supérieur, Université de Constantine, Faculté des Sciences de l'ingénieur, Département Informatique (2004)
5. Maedche, A., Zacharias, V.: Clustering Ontology-based Metadata in the Semantic Web: Principles of Data Mining and Knowledge Discovery. In: 6th European Conference, PKDD, Helsinki, Finland (2002)
6. Occupationnel information network (O*NET), <http://online.onetcenter.org/>
7. Schmidt, A., Kunzmann, C.: Professional Learning Ontology and Competencies. In: the OntoContent workshop at the On the Move Federated Conferences, Montpellier (2006)
8. Trichet F., Radevski V.: Competency-based Systems dedicated to e-Recruitment. In: Proceedings of the 4th International Conference on Computer Science and Information Technology. pp 185-197, Volume3, ISBN: 9957-8592-0-X, Amman (CSIT'2006)
9. Yahiaoui, L., Boufaïda, Z., Prié, Y.: Automatisation du recrutement dans le cadre du web sémantique. In: 17 journées francophones d'ingénierie des connaissances, Nantes, (IC 2006)
10. Yahiaoui, L., Boufaïda, Z., Prié, Y.: Semantic annotation of documents Applied to E-Recruitment. In: Proceedings of SWAP, the 3rd Italian Semantic Web Workshop, Pisa, Italy, (2006)