IADIS International Journal on Computer Science and Information Systems
Vol. 6, No.1, pp. 30 - 44
ISSN: 1646-3692

USING TAXONOMIES AND ONTOLOGIES TO DEFINE OCCUPATION AND EDUCATION-DRIVEN EUROPEAN QUALIFICATIONS

Valentina Gatteschi. *Dipartimento di Automatica e Informatica – Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129, Torino – Italy.
valentina.gatteschi@polito.it

Fabrizio Lamberti. *Dipartimento di Automatica e Informatica – Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129, Torino – Italy.
fabrizio.lamberti@polito.it

Andrea Sanna. *Dipartimento di Automatica e Informatica – Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129, Torino – Italy.
andrea.sanna@polito.it

Claudio Demartini. *Dipartimento di Automatica e Informatica – Politecnico di Torino*
Corso Duca degli Abruzzi 24, 10129, Torino – Italy.
claudio.demartini@polito.it

ABSTRACT

The mobility of students and workers is a relevant topic in the Vocational Education and Training field. However, despite the introduction of several European instruments, like the European Qualification Framework and the European Credit system for Vocational Education and Training, there are still several barriers to the recognition of qualifications acquired abroad. The above limitation could be overcame by the creation of a European common profile that could be used by each education and training Authority as a reference for the corresponding national qualifications. In this paper, a methodology exploiting an ontology and a taxonomy-based approach to identify a common profile within the European trade sector is presented. The proposed methodology allows the users to compare the requirements of the labor world and the outcomes of the education and training routes with the aim of finding a common denominator. Existing qualifications could then be added with elements belonging to the common profile, in order to make them transparently recognized in a true transnational perspective.

KEYWORDS

Taxonomy, ontology, lifelong learning, qualification comparison, semantic engine.

1. INTRODUCTION

Lifelong learning and mobility for learning are taking place increasingly in a wide variety of contexts; formal, non-formal and informal and in different countries. Mobility can be an instrument to address existing skill or labor market shortages and skills mismatches in a country or region, thus improving the efficiency of labor markets and removing brakes on economic growth.

The question of course is, if all parties involved do in principle benefit from transnational mobility, why – on a proportional basis – does not the participation in transnational mobility by young people in vocational training increase substantially? It seems that transparency of qualifications, and moreover the differences in the meaning, content and interpretation of tasks and functions on the European labor market and the mutual agreements on qualification profiles are one of the barriers in the mobility of learners and workers.

In order to overcome this limitation, in recent years, European legislators made a huge effort for improving students' and workers' mobility: in fact, several initiatives have been accomplished in order to guarantee comparability, transferability and recognition of qualifications across different countries, as well as to enhance transparency and mutual understanding across Member States, with the aim of overcoming the gap between heterogeneous education and training systems and ensuring that "the European labor market is open to all", as it is expected from the Bruges-Copenhagen process.

Nevertheless, even if Higher Education (HE) already presents qualifications depicted according to standard rules, Vocational Education and Training (VET) training paths are still lacking a uniform description, thus limiting students transfers between countries.

A first attempt to create a shared understanding in the lifelong learning domain has been done in 2008, by the European Parliament Council, which established the European Qualification Framework (EQF) [12], a common reference system conceived to support the linking of different countries' national qualifications systems and frameworks together. According to the EQF, this could be done by exploiting a rigorous classification of lifelong learning qualifications based on eight reference levels, and by defining precisely the semantics of associated learning outcomes, in terms of knowledge (*the body of facts, principles, theories and practices that is related to a field of work or study*), skills (*the ability to apply knowledge and use know-how to complete tasks and solve problems*) and competence (*the demonstrated ability to use knowledge, skills and personal, social and/or methodological abilities, in work or study situations and in professional and personal development in order to achieve objective results according to a specific level of autonomy and context complexity*) concepts, thus opening the way for the creation of a shared understanding in the lifelong learning domain.

Even though the definition of a European-wide framework could provide a shared format for the description of qualifications, other tool – such as instruments for supporting students and workers who want to continue their training or working career abroad, or companies who are looking for workers with specific abilities – have to be developed in order to increase mobility. Consequently, it is clearly visible that, in order to be able to work at a transnational level, the above instruments should rely on descriptions of qualification (and composing learning outcomes) achieved by a student or worker, based on a standard and syntax-independent formalism, i.e. by exploiting strategies and tools developed in Semantic Web initiatives.

In this paper we give an overview of the results achieved by the TIPTOE project, a transnational project funded under the Lifelong Learning Programme (LLP) that aimed at exploiting a semantic platform for the construction of a common European profile in the trade sector. Specifically, the designed environment presented in this work exploits a semantic engine that is able to perform an EQF aware taxonomy-based comparison of country-based formative offers expressed through educational profiles and labor market requirements represented by national occupational profiles with the goal of finding similarities and specificities emerging from heterogeneous "local" descriptions structured in terms of knowledge, skills and competences.

The rest of paper is organized as follows: Section 2 analyses and discusses approaches for describing qualifications and for performing automatic processing onto them defined in previous research activities. In Section 3, the main idea behind the present work is explained, by making reference to the four-stage methodology developed within the project, encompassing data collection, formalization, semantic reasoning and production of result. Section 4 presents an example of application of the above methodology in the trade sector and reports experimental achievements. Finally, conclusions and future developments are developed in Section 5.

2. RELATED WORKS

As briefly presented previously, this work aims at identifying common elements or differences between European occupational and training worlds, by exploiting core elements belonging to the Semantic Web field. In particular, core concepts such as ontologies (defined by Gruber as *explicit specifications of a conceptualization* [6]) have been utilized as a support for the comparison and linking of qualifications, whereas a taxonomy (a classification arranged in a hierarchical structure) has been created in order to define relations among concepts and, thus, to improve the comparison and linking processes.

The application of the semantic paradigm to working and learning domains has already been investigated in the literature; a first example is represented by the CUBER-project [10], where a system exploiting standardized metadata in order to support learners in searching higher education courses that match their needs is presented. A different approach, that goes beyond the solution in [10] is shown in [8]: in this work, a semantic search strategy based on the analysis of the relations among concepts belonging to user queries and concepts used in learning documents is presented. A different strategy aimed at increasing students' mobility is depicted in [1]; here, the author suggests an interesting use of taxonomies for comparing European engineering courses: in particular, an adaptation of the Bloom's learning taxonomy for organizing verbs is investigated. A further approach that defines a general-purpose strategy for measuring the differences among qualifications by defining meta-ontologies describing referencing rules between national models is presented in [5]. In this work, formal models of national education and training systems are created, and meta-ontologies are exploited for overcoming the heterogeneity of qualification structures belonging to different countries. Another interesting methodology is presented in [2]: in this work, the authors try to reduce the cognitive overload and the cognitive disorientation that may arise by the inappropriate curriculum sequencing during personalized Web-based learning, and present a novel genetic-based curriculum sequencing scheme based on a generated ontology-based concept map. A further approach, that shows how ontologies could be exploited in order to enhance and maintain curricula belonging to students of medical science is shown in [3]: this research presents an ontology-based knowledgebase able to deal with the multi-dimensional matrix depicting relations between curriculum contents and intended learning outcomes. The issue of curriculum development has been tackled also in [11], where the authors define an ontology of electrical engineering curricula that takes into account correlated topics that might exist in various courses, and allows the system to identify, when a new course is added, which are the pre-requisites and the foundational courses to be added to the study plan. Another work that also deals with the need for a personalized curriculum in the European perspective is presented in [4]: in this research, the authors present a common conceptual model supported by the Academic Ontology Bologna Process in order to allow interoperability between academic management systems and automation of academic management. Finally, in [7] the use of a domain ontology for automatically producing a semantic annotated electronic résumé is proposed: according to the authors, the recruitment phases could be possibly supported by an ontology of terms, that could be used to suggest, starting from an initial set of competences specified by the user, additional competences that may be also included in the résumé.

As in some of the above works, the methodology discussed in this paper exploits taxonomies and ontologies for representing qualifications and for developing semantic-based comparison strategies. However, while the above works generally aim at ranking elements according to their degree of similarity with a target description, the objective of the present work is the identification of common elements among a huge variety of descriptions. Moreover, the current work proposes a general-purpose methodology strongly exploiting subsumption relations within a strictly structured context represented by the EQF framework and by its associated principles, and aims at investigating whether the use of classifications of concepts could improve (or worsen) the results of the comparison.

3. PROPOSED METHODOLOGY

The education and training domain, and in particular the VET scenario, is characterized by a marked difficulty to find qualifications recognized at the European level. In fact, qualification descriptions and syllabi of VET courses are frequently drawn at the national (or regional) level, or even by the training body responsible for the course. The lack for rules outlining a minimum set of knowledge, skills and competences that a student should possess after a training path creates strong information asymmetries between the education and the labor worlds, and severely limits the mobility among countries.

In fact, since a unique and well-defined qualification profile is missing, employers may ignore the exact contents of the courses attended by a student who is applying for a given job position, and consequently, may not know which knowledge, skills and competences he or she actually achieved. The depicted scenario is even more jeopardized and complex in a transnational perspective, especially when non-formal and informal learning paths are also taken into account.

A possible approach to overcome the above barriers is presented in this work, where the results of the research activities carried out in the frame of a transnational project aimed at defining a methodology for creating a unique European profile and, finally, at applying it in the specific domain of the trade sector are summarized. It is worth remarking that, although in the TIPTOE project the devised approach was applied to a specific economic sector, the proposed methodology could be considered as totally general-purpose; hence, its exploitation in other contexts should be straightforward.

The strategy behind the present work is the following: first, an investigation by means of a series of interviews with relevant stakeholders is performed on the labor world. This first phase aims at outlining a set of tasks a worker should be able to accomplish by characterizing them in terms of knowledge, skills and competences. Secondly, the education and training field is investigated: in this phase, several interviews with relevant training organisms are conducted in order to identify which knowledge, skills and competences a student should possess at the end of a formal training path.

Once collected, the two sets of information must be compared in order to find the common elements between the requirements of the occupational domain and the outputs of the educational routes: this comparison is oriented to the definition of a unique profile, drawn according to the EQF principles. When considering the huge amount of data to be evaluated, the comparison operation would risk to be extremely time-consuming and to provide incorrect results, if carried out manually.

As a consequence, suitable instruments should be exploited, allowing to catalogue the outcomes of the interviews in a structured way and to perform the required semantic reasoning onto them. The modus operandi pursued in the TIPTOE project consisted in linking the elements belonging to occupational and educational descriptions to a set of concepts, organized into a taxonomy: the just drawn ontology – which shows the links among elements of the descriptions and concepts of the taxonomy – could then allow to carry out the necessary reasoning by exploiting the relations among elements and, thus, overcoming lexical barriers.

As anticipated, the above methodology consists of four stages: *information collection*, *taxonomy and ontology construction*, *definition of inference rules and approaches for semantic comparison* and, finally, *common profile creation and analysis of results*.

3.1 Information Collection

As already mentioned, the *information collection* phase is aimed at collecting the requirements of the labor world and the outputs of the education and training domain, expressed in terms of task and subtasks as well as of knowledge, skill and competence elements.

It is worth remarking that, in order to define a shared format for collecting information (and then representing it in the taxonomy construction stage) the representation of knowledge, skill and competence concepts made by [9] was exploited. According to [9], a knowledge could be seen as a set of knowledge objects (KO). A skill could be defined as a KO “put into action” through an action verb (AV), hence by one or more pairs KO – AV. Finally, a competence could be represented as a triple KO – AV – CX, that describes the ability of putting into action a given KO in a specific context (CX). The information collection stage was carried out by keeping in mind the relations among the above concepts.

Within the TIPTOE project, the scheduled interviews were performed in order to investigate the relevant elements of four professional profiles in the selected sector, namely Shop Assistant, Shop Manager, Logistic Assistant and Logistic Manager.

At a first stage, stakeholders (i.e. employers of the retail and wholesale sectors) belonging to the labor context of different European countries (i.e. France, United Kingdom, the Netherlands, Italy, Lithuania, Portugal, Germany and Slovenia) were interviewed in order to collect, for each of the four profiles above, a list of knowledge, skills and competences that a worker must possess for fulfilling a task, each characterized by the corresponding EQF level (depicting the complexity degree). After the first interview phase, the education domain was investigated by interviewing vocational training Authorities of the eight countries above, with the aim of collecting information regarding learning outcomes achieved by the students at the end of a specific training route.

After having fulfilled this phase, several grids were filled in with information concerning tasks, subtasks, knowledge, skills and competences.

3.2 Taxonomy and Ontology Construction

Profiles collected in the previous phase were then inspected in order to identify the core elements to be structured in a taxonomy (knowledge objects, action verbs and context elements). In particular, for each instance of knowledge, skill and competence elements, one or more concepts (or keywords) were selected.

Concepts were then linked to each other by exploiting subsumption relations, and the whole set of elements was expressed according to a taxonomic representation. Three families of concepts were created in the hierarchical tree: a first one made up of knowledge object, a second one including action verbs and a third one depicting the context.

For what it concerns the creation of the knowledge and the context families, it was necessary to start from scratch, since the existing taxonomies were not able to fully satisfy our requirements. On the contrary, action verbs were initially structured by exploiting the Bloom's taxonomy [1] and by adapting it: at the end of this process, six families of verbs, *arrange*, *act*, *prepare*, *check*, *assess* and *react* were identified.

It is worth remarking that the definition of the taxonomy was a crucial point, since an improper hierarchy of keywords could provide incorrect results: thus, this phase was performed with the support of experts from the trade sector.

Then, after the creation of the taxonomy, qualifications, tasks and subtasks were described by linking their composing elements (knowledge, skills and competences) to the corresponding concepts (knowledge objects, action verbs and context elements).

During this phase, a graphical representation of the ontology was drawn by exploiting the UML notation, in order to provide a formal and easy-to-read description to be possibly shared with the other involved actors and stakeholders. For this purpose, the open-source tool UMLGraph [13], a software that is able to process diagrams expressed in a textual form and to draw the corresponding graphical representation was exploited. It is worth remarking that UMLGraph was selected since it could easily be embedded inside a web platform as a tool for the construction of the profile maps requested by the user. Figure 1 depicts an excerpt of the subtask *To welcome the customer and understand the customer's needs and requests*, belonging to the Portuguese Shop Assistant profile: more specifically, the diagram displays the knowledge *Communication techniques knowledge*, the two skills *To be able to apply* *selling techniques* and *To be able to communicate in English*, and the competence *Full responsibility in identifying the customer and his needs*. In order to better characterize knowledge, skill and competence elements, the corresponding classes are colored black, while the concepts of the taxonomy they are linked to are painted light gray.

Moreover, subsumption relations are expressed by a solid line with a hollow arrowhead that points from the class that is subsumed to the class that subsumes. Finally, the fact that a knowledge, skill or competence is characterized by one or more concepts belonging to the taxonomy is indicated by a dashed line. It is worth remarking that this type of lines has been used in order to make more readable the diagram, so that it is immediately understandable which are the relations that define subsumption of terms, and which are the relations that link knowledge, skills and competences to the taxonomy (in this case a dashed line is drawn for showing the link between an element of a subtask and a concept in the taxonomy, or among knowledge elements, action verbs and context, and not for indicating dependency relations).

The diagram should be interpreted as follows: the knowledge *Communication techniques knowledge* is characterized by the knowledge object *communication techniques* that is a type of *selling techniques*, i.e. another knowledge object. The subsumption relation between *selling techniques* and *communication techniques* indicates that if someone has got a *communication techniques* knowledge, he or she has also a *selling techniques* knowledge. Furthermore, the skill *To be able to apply selling techniques* is described by the pair of concepts *apply*, an action verb that specifies the action verb *act*, and *selling techniques*, a knowledge object, while the skill *To be able to communicate in English* is defined by the action verb *communicate*, a specification of the action verb *react*, that is applied to the *English* concept, a specification of a *foreign language* knowledge object. Finally, the competence *Full responsibility in identifying the customer and his needs* is characterized by a *full responsibility* context, applied to the *identify* action verb, that is linked to *customer* and *customer needs* knowledge objects.

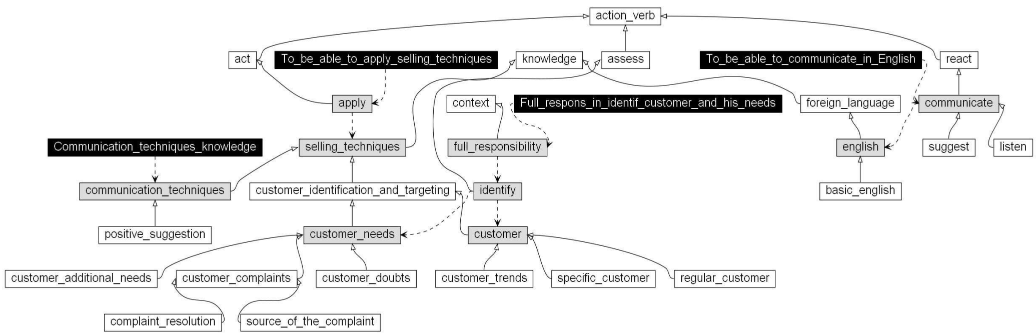

Figure 1. Ontology related to the subtask: *To welcome the customer and understand the customer's needs and requests*

3.3 Definition of Inference Rules and Approaches for Semantic Comparison

The hypothesis at the basis of the definition of the inference rules required for the TIPTOE project is the following: since the common profile has to act as common denominator, necessarily it should be a combination of elements that are present into all the profiles and, as a consequence, it should be the sum of all the knowledge, skills and competences that are linked to the most used knowledge objects, action verbs or context elements.

The above statement could probably be explained in a clearer way by the following example: let us imagine that four subtasks belonging to four profiles are defined by the following knowledge: *cleaning techniques knowledge*, *cleaning means and tools knowledge*, *cleaning methods knowledge* and *cleaning methods*, *means and tools*; since each profile contains (at least) a knowledge that is related to the cleaning activity, this knowledge should appear also in the common profile. On the contrary, if a knowledge, e.g. the *product lifecycle knowledge*, is mentioned in only one profile, it will not have be incorporated into the common profile.

Moreover, the reasoning that has been described above should be based on the keywords linked to the elements belonging to the profile descriptions. In fact, a semantic engine should be able to understand that the four knowledge elements mentioned above are linked to the cleaning concept (then, in the ontology, they will be characterized by the *cleaning knowledge* element).

It is evident that the common profile will then be the sum of the most common knowledge, skill and competence elements. As a consequence, the engine for semantic comparison should be able to find the most used keywords, recognize to which elements they are linked, and then include these elements into the common profile. A further step towards the achievement of a more correct result could be the exploitation of the taxonomy of terms and subsumption relations: in this way, by examining the example shown in Figure 1, the number of occurrences of *communication techniques*, *customer*, *customer needs*, *English*, *apply*, *identify*, *communicate* and *full responsibility* would be 1, while the number of occurrences of the (parent) element *selling techniques* would be 4, since the *selling techniques* concept has been exploited once, but the (children) elements *communication techniques*, *customer* and *customer needs* have been used each one once too.

In order to find the best result, four comparison strategies have been developed; all of them take as input a threshold (that is a minimum number of times a keyword has to be used) defined by the user, and explore the ontology in order to identify the most common elements.

The four comparison strategies developed are: *simple range*, *simple range with mean*, *aggregate simple range* and *aggregate range with mean*.

The *simple range* strategy is the simplest way of determining which knowledge, skill and competence elements will belong to the common profile, since it calculates a value that corresponds to the number of times a keyword has been linked to the elements of the ontology; if this value is higher than the threshold defined by the user, the strategy includes the considered knowledge, skill or competence into the common profile.

A slightly more complex approach is the *simple range with mean*: according to this strategy, the value computed by the comparison tool (which, in order to add the element belonging to the common profile, must be higher than the threshold defined by the user) is the average of the number of occurrences of each keyword linked to the knowledge, skill or competence being considered.

A third approach, which takes into account also hierarchical relations expressed by the taxonomy, is the *aggregate simple range*: according to this strategy, the tool calculates the number of times a keyword, and the subsumed concepts, have been used to describe the elements of the ontology; if this value is higher than the threshold specified by the user, the examined element is added to the common profile.

A fourth strategy, that is similar to the *simple range with mean* and that allows to consider also subsumption, is the *aggregate range with mean*: according to this approach, the value computed by the comparison tool is the mean of the number of occurrences of each keyword and its children in the taxonomy.

In order to better understand the logic behind the four different approaches, it could be useful considering a further example, e.g. represented by the *Knowledge of products and relevant display techniques (i.e. volume displays and on shelf couponing)* element: let us assume that this knowledge is described by the keywords *product* (used 38 times in the profile descriptions), *exposition techniques* (used 12 times), *volume displays* (used 3 times) and *on shelf couponing* (used only in this description). Furthermore, let us suppose that the *product* and the *exposition* elements have several children in the knowledge taxonomy, and that the respectively subsumed classes have been used 84 times and 14 times, respectively.

When the simple range approach is followed, the result is 38, that is the maximum value of occurrences of the keywords linked to the knowledge. On the other hand, the result of the simple range with mean strategy is 13.5, that is the average of the occurrences of the four keywords linked to the element.

When subsumption relations are taken into account, the computed value increases: in fact, according to the aggregate range approach, the result is 122, that is the sum of the occurrences of *product* (122, that is 38+84), *exposition* (26, that is 12+14), *volume displays* (3) and *on shelf couponing* (1) concepts, whereas if the strategy applied is the aggregate range with mean, the result is 38, that is the average of the values above.

It is worth remarking that the results just explained (and shown in Table 1) are only an estimate of how common a knowledge, skill or competence is; hence, a given value could not be good or worst a priori, because it has to be compared with the other results. As a consequence, possible ways for identifying the common profile could be to order the results from the one that obtained the highest value, to the worst one, and then select a number of elements defined by the user (i.e., the number of knowledge, skill and competence elements in the common profile would be fixed), or – and this is the case – to use the threshold expressed by the user to select only those elements that achieved a score higher than it.

Table 1. Results obtained from the application of the comparison strategies to the *Knowledge of products and relevant display techniques (i.e. volume displays and on shelf couponing)* element

Strategy	Result
Simple range	38
Simple range with mean	13.5 = (38 + 12 + 3 + 1) / 4
Aggregate range	122 = (38 + 84) + (12 + 14) + 3 + 1
Aggregate range with mean	38 = [(38 + 84) + (12 + 14) + 3 + 1] / 4

3.4 Common Profile Creation

The approaches explained in the previous stage were exploited for the creation of the common profile. According to the above discussion, the knowledge, skill and competence elements obtaining a specific value become potential components of the common profile. However, since – as in the case of the above example with the set of knowledge described by the cleaning techniques concept – it would be redundant inserting into the common profile four elements with the same meaning, we decided to let the user choose, among the set of elements exploiting the same keywords, the one that could better represent the specific knowledge, skill or competence.

An example is shown in Figure 2 where, concerning the knowledge described by the knowledge object *cleaning techniques*, the user selected as the representative knowledge the element *cleaning methods*; moreover, among the elements linked to *communication techniques knowledge*, the user selected *knowledge of selling and communication techniques*.

Finally, for each common profile an EQF level was computed as an average of the EQF values assigned to each knowledge, skill or competence selected as potential components of the whole profile.

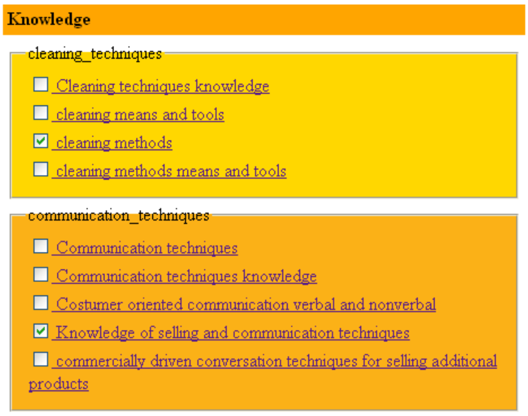

Figure 2. Selection of the elements that will belong to the common profile

4. EXPERIMENTAL RESULTS

The depicted methodology has been followed for the construction of a common European profile for each one of the four trade qualifications considered in the study. In particular, for each profile, eight descriptions of job positions in the labor market field and qualifications in the educational domain have been collected.

In the following, experimental results concerning the Shop Assistant profile will be analyzed. For this job position, the total number of learning outcomes inserted and organized in the ontology, together with an EQF level identified by the experts, has been 1096 for labor market and 1095 for the training dimension.

In order to perform a quick analysis of the inserted profile, the wordle tool [14] has been exploited for the creation of a tag cloud of the composing learning outcomes (Figure 3).

Figure 3. Tag cloud of the learning outcomes related to the Shop Assistant profile

Even though this tool allows users to identify a profile by giving a quick look to the tag cloud (since the tag cloud of a managerial job positions will be a combination of concepts denoting higher responsibility, while an assistant job visual description will be characterized by more lower-level activities), the size of a word is only linked to the number of its occurrences in a sentence, without taking into account also relations among terms. An example is shown in Figure 3: in this representation, terms like *goods*, *product*, (product) *characteristics* are not related to each other, while, in reality they refer to the same thing (the *product*); as a consequence, the number of occurrences of this group of terms, instead of the value expressing single occurrences, should be considered. Hence, composing concepts have been organized in a taxonomy, consisting, for the Shop Assistant qualification, of around 600 elements.

Figure 4 shows a statistic of the most used concepts in the description of learning outcomes belonging to the Shop Assistant job profile: in the chart, the occurrence of single concepts and aggregate concepts (computed by analyzing subsumption relations) is depicted. An interesting behavior is related to the *selling technique* concept. In fact, this knowledge element is certainly important for a Shop Assistant; however, only few learning outcomes, if compared to the *product* knowledge element, mentioned it. Hence, a strategy counting single occurrences (blue bars) would probably not be able to identify it as a common element. On the contrary, when aggregate values are exploited for the creation of the European profile (orange bars), this knowledge acquires more importance, by nearly becoming as relevant as the *product* knowledge.

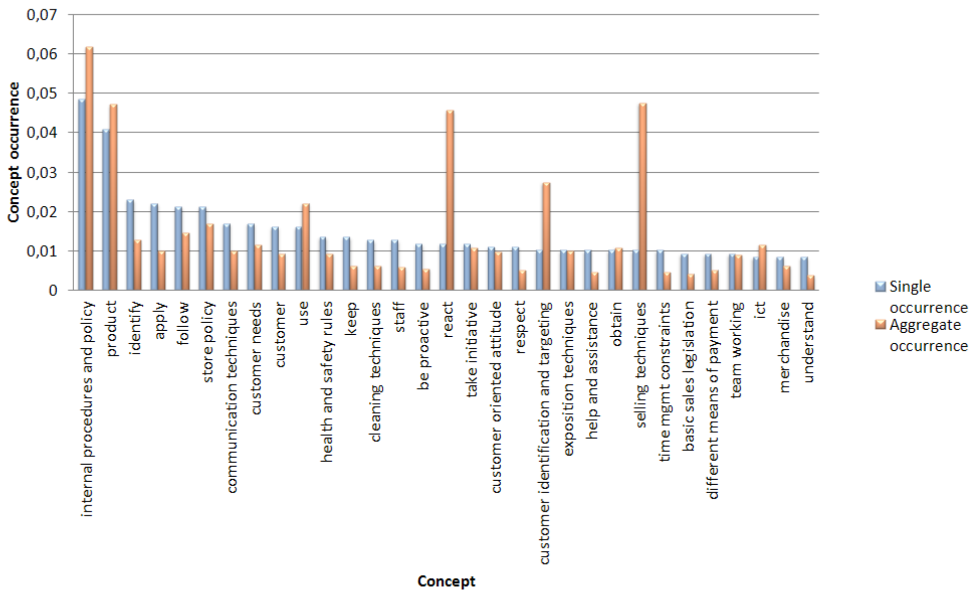

Figure 4. Most used concepts in the description of learning outcomes characterizing the Shop Assistant job profile

Having considered all of the above, we could say that, in general, "simple" approaches (*simple range* and *aggregate range*) perform better with long sentences characterized by the description of a knowledge and several examples, like *Knowledge of products and relevant display techniques (i.e. volume displays and on shelf couponing)*; but they provide worst results with skills and competences defined by a common verb and an uncommon noun (an example could be the skill *to apply stocktaking procedures*, since this element obtains a high rate, even if the stocktaking procedure knowledge is a rare concept). On the contrary, a strategy that computes the average of the occurrences produces better results with elements like the skill above, but risks to provide worst results in case of long sentences.

For what it concerns the exploitation of subsumption relations, if an approach that does not consider the taxonomy (*simple range* and *simple range with mean*) is pursued, all the keywords (children and parents) have the same importance but, if the different profiles are described with a huge variety of terms, incorrect results could be achieved; a strategy that considers also the taxonomy (*aggregate range* and *aggregate range with mean*) implicitly interprets as more important the highest elements of the tree (the parents) and, for this, it could be useful to overcome lexical differences.

Clearly, it is impossible to identify a priori the best approach, since the adoption of a particular strategy would depend on the description of the elements of the ontology: the choice of the approach to be followed could only be done a posteriori. However, taking into account the linguistic barriers we encountered during the previous stages, we identified as best approaches for creating a common profile the *aggregate range with mean* and the *aggregate range*.

Table 2 shows the learning outcomes that should be included in a common European profile, identified by means of the *aggregate range* strategy. With respect to the EQF level, we experienced that, when experts had to assign an EQF level to a profile, they first tried to identify the level of each learning outcome (by studying the degree of responsibility and autonomy, the characteristics of the knowledge elements, etc.), then they determined the overall level by finding the value occurring the most. Hence, we decided to compute the EQF level of each learning outcome that will appear in the common profile as the average of the value of learning outcomes described by the same concepts (where, indeed, fractions have to be interpreted by making reference to the closest integer value). In this view, the EQF level of the skill *To be able to describe clearly the products and theirs specificities* denotes that the majority of the learning outcomes described by the concepts related to the *product specificities* knowledge and the *describe* action verb are characterized by an EQF level three (or higher), and only few skills are assigned an EQF level two (or lower).

Table 2. Results of the *aggregate range* strategy for the Shop Assistant job profile

Knowledge	EQF	Skills	EQF	Competence	EQF
Costumer oriented communication verbal and nonverbal	2.00	To be able to apply proper communication technique	2.40	Responsible for an effective and appropriate welcoming for the identification of customer requests and needs and for identifying how many products can that specific customer buy	3.00
ICT knowledge	2.40	To be able to apply selling techniques	2.50		
Internal procedures knowledge	2.13	To be able to communicate in English with the customer	3.00		
Knowledge of cash register operations, invoicing, means of payment, and VAT refund form	2.75	To be able to describe clearly the products and their specificities	2.90		
Knowledge of processes for providing customer service and assessing customer needs	3.00	To be able to differ customer types	2.00		
Knowledge of products	3.00	To be able to follow internal procedures	2.75		
Knowledge of selling and communication techniques	2.50	To be able to handle cash and other payment means	2.00		
Knowledge of store policy and basic sales legislation	4.00	To be able to identify shop logistic needs	2.86		
Knowledge of organization and arrangement of merchandise	2.00	To be able to understand which products are more adequate to customers s needs	2.42		
Knowledge of differences between specific and regular customers	2.00	To be able to use and take care of tools for goods movement in accordance with the legal regulations	3.00		
		To be able to use the Internal Registration program to register all products delivered	2.41		
		To be able to work with ICT	3.00		
		To be able to act on quantitative and qualitative deviations	2.00		
		To be able to apply company specific procedures in receiving and approaching customers	2.00		
		To be able to keep a customer oriented attitude react quickly be proactive take initiatives	3.14		

5. CONCLUSION AND FUTURE WORKS

In this paper, a methodology for the comparison of occupational and educational profiles and for the identification of their common elements is presented. According to the proposed approach, once a profile, defined in terms of a set of task and subtasks – that require several knowledge, skill and competence elements – has been expressed as a set of knowledge objects, action verbs and context components (further organized into a taxonomy), then it becomes possible to perform a semantic analysis and to find which elements appear more frequently in the profile descriptions. In order to identify the best approach for the definition of a common profile, four strategies for the comparison are defined.

The proposed approach could also be exploited for supporting the definition of new syllabi, since it defines guidelines for the comparison of the requirements of the labor world and the outputs of the education and training domain. Future works will be addressed to extend the devised methodology by implementing algorithms exploiting subsumption relations to identify the differences and the missing elements among profiles.

ACKNOWLEDGEMENT

This work was partially supported by the TIPTOE "Testing and Implementing EQF and ECVET Principles in Trade Organizations and Education" project funded under the Lifelong Learning Programme - Leonardo da Vinci (NL08LLP-LdVTOI123011).

REFERENCES

[1] Bloom, B.S. et al, 1956. *Taxonomy of educational objectives – the classification of educational goals – Handbook I: cognitive domain*, Longmans, USA.

[2] Chih-Ming Chen, Chi-Joi Peng and Jer-Yeu Shiue, 2008. Ontology-based Concept Map for Planning Personalized Learning Path. *IEEE Conference on Cybernetics and Intelligent Systems (ICCIS 2008)*, pp.1337-1342.

[3] Dexter, H. and Davies, I., 2009. An ontology-based curriculum knowledgebase for managing complexity and change. 9th *IEEE International Conference on Advanced Learning Technologies (ICALT 2009)*, pp.136-140.

[4] Ferreira, J.C. and Filipe, P., 2009. Academic Ontology to Support the Bologna Mobility Process. *2nd International Conference on Adaptive Science & Technology (ICAST 2009)*, pp.308-313.

[5] Gatteschi, V. et al, 2009. An Automatic Tool Supporting Life-long Learning based on a Semantic-oriented Approach for Comparing Qualifications. *IADIS International Conference on Cognition and Exploratory Learning in Digital Age (CELDA 2009).* Rome, Italy, pp 354-358.

[6] Gruber, T.R., 1993. Padua workshop on Formal Ontology. *Toward principles for the design of ontologies used for knowledge sharing*, pp 1-23.

[7] Mirizzi, R. et al, 2009. A Semantic Web enabled System for Résumé Composition and Publication. *Semantic Web Information Management (SWIM 2009).* pp 583-588

[8] Nemirovskij, G. et al, 2007. SWAPS: Semantic Web Approach for Personalisation of Study. *7th IEEE International Conference on Advanced Learning Technologies (ICALT 2007)*, pp. 711-712.

[9] Pernici, B. et al., 2006. The eCCO System: An eCompetence Management Tool Based on Semantic Networks. *Workshop on Ontology Content and Evaluation in Enterprise (OnToContent 2006)*, pp. 1088-1099.

[10] Pöyry, P. and Puustjärvi, J., 2003. CUBER: A Personalised Curriculum Builder. *3rd IEEE International Conference on Advanced Learning Technologies (ICALT 2003)*, pp.326-327.

[11] Tang, A., 2009. An Ontological Approach to Curriculum Development. *International Conference on Engineering Education (ICEED 2009)*, pp.219-224.

[12] The European Qualifications Framework for Lifelong Learning (EQF) on-line: http://ec.europa.eu/education/ lifelong-learning-policy/doc44_en.htm (February 17, 2011).

[13] UMLGraph on-line: http://www.umlgraph.org/ (February 17, 2011).

[14] Wordle on-line: http://www.wordle.net/ (February 17, 2011).