Article

Lifelong Learning Courses Recommendation System to Improve Professional Skills Using Ontology and Machine Learning

María Cora Urdaneta-Ponte [ID], Amaia Méndez-Zorrilla *[ID] and Ibon Oleagordia-Ruiz [ID]

eVIDA Research Group, Faculty of Engineering, University of Deusto, 48007 Bilbao, Spain;
cora.urdaneta@deusto.es (M.C.U.-P.); ibruiz@deusto.es (I.O.-R.)
* Correspondence: amaia.mendez@deusto.es

Featured Application: This article describes a system for recommending lifelong learning courses for the improvement of professional skills by using an ontology that is automatically updated to model professional profiles and courses.

Abstract: Lifelong learning enables professionals to update their skills to face challenges in their changing work environments. In view of the wide range of courses on offer, it is important for professionals to have recommendation systems that can link them to suitable courses. Based on this premise and on our previous research, this paper proposes the use of ontology to model job sectors and areas of knowledge, and to represent professional skills that can be automatically updated using the profiled data and machine learning for clustering entities. A three-stage hybrid system is proposed for the recommendation process: semantic filtering, content filtering and heuristics. The proposed system was evaluated with a set of more than 100 user profiles that were used in a previous version of the proposed recommendation system, which allowed the two systems to be compared. The proposed recommender showed 15% improvement when using ontology and clustering with DBSCAN in recall and serendipity metrics, and a six-point increase in harmonic mean over the stored data-based recommender system.

Keywords: lifelong learning courses; ontology; machine learning; hybrid system recommendation

1. Introduction

Organisations are currently undergoing significant changes as evidenced by the continuous and rapid transformation in response to globalisation and to advances in information and communication technologies. The resulting changes in job requirements mean that professionals must improve and/or develop competencies and skills in line with job market trends in order to make them successful and ensure that businesses are competitive.

According to study [1], employability requirements have changed to address these new situations and challenges, and discrepancies have emerged in some areas between the employees' skills and labour market needs. This has revealed that the knowledge acquired by graduates in a higher education course of study is outdated by the time they obtain their diplomas.

Employers today focus on attracting good employees who often not only have basic academic skills but also higher-order thinking skills such as learning, reasoning, creative thinking, decision making and problem solving [2]. The same authors also noted that a study by the Stanford Research Institute and the Carnegie Melon Foundation found that 75 percent of long-term job success depends on people skills, and 25 percent on technical knowledge. Organisations place importance on employability skills, which are those needed to get, keep and do a job well [2], and can be defined as a range of skills

or competences that can be developed over a lifetime through education, training, work experience, interests and extracurricular activities.

According to [3], these changes have been reflected in some elements of the education system. Increased global economic competition has led to an emphasis on lifelong learning with a focus on market-based skill building [4]. The market has moved from demanding professionals with skills adapted to an industrial mode of production to needing professionals who must develop skills that are more closely related to current economic and social developments [5].

Study [6] indicated that lifelong learning is the process of expanding knowledge and procedures, professional skills and attitudes using deliberate, self-directed efforts. According to [7] lifelong learning is learning that occurs during all stages and spheres of life. The aim of lifelong learning is to efficiently improve performance and to be able to adapt successfully to emerging short-and long-term challenges and new requirements throughout the career lifespan.

The need to have professionals with up-to-date skills has led to a vast offering in lifelong learning. There is a very wide range of courses available to cover different areas of knowledge, which are continuously updated in order to adapt to constant changes and ensure that individuals are fit for professional practice. A set of skills and abilities are thus obtained that can be described as know-how, resulting from the integration and adaptation of skills and abilities to typical situations [8] necessary to perform effectively in the labour market.

The growth of the internet in recent years has expanded the number of courses that can be found on the web. Despite users having more information available, it has become increasingly difficult for them to find courses that meet their demands according to their needs and preferences. The use of Social Networking Sites (SNSs) to establish professional links and exchange information is becoming increasingly common [9]. SNSs have become an important source of information for both the creation and dissemination of content, including a large amount of information related to professional profiles, job offers and even learning opportunities. According to a study by [10], the use of SNSs has led many organisations to incorporate this new source of information into their selection processes, Facebook and LinkedIn being the most widely used. This is due to the type of information in the work domain they share, such as job, education and skills; characteristics to take into account when proposing a learning pathway to improve and/or develop professional skills.

Users need to prioritise filtered data to efficiently access the right information on the web. In the era of Big Data, it is challenging to find relationships and patterns from within vast amounts of mostly unstructured or semi-structured information. In this context, an important role is played by recommendation systems (RSs), which are mainly aimed at predicting user ratings or preferences of an item that are valuable to the user [11]. Examples of popular recommendation systems include Amazon, especially in the commerce sector; Facebook, that can recommend social connections and Netflix, which offers movie recommendations based on user preferences and viewing history [12].

RSs handle a large amount of data which, depending on their nature, may come from different sources and are characterised by their heterogeneity. The quality of recommendations largely depends on the retrieval and representation of information. In this process there may be problems related to the diversity of types and sources of data used, or to the lack of information provided by users or recommended elements, known as a cold start [12], which can be a challenge for this type of system. The theory of semantic knowledge representation has been adopted to meet this challenge and help solve these problems [13].

In order to represent reality in a specific domain, so that it is readable by machines, the semantic web uses ontologies [14]. An ontology is a formal specification of the types, properties and interrelationships of entities for a particular domain. Since ontologies are generally built on the basis of domain expert consensus, they are highly reliable structured knowledge bases with a wide range of applications [15]. Some ontology-based RSs have

been developed which allow information to be extracted and knowledge about users, items and the relationship between them to be represented in the application domain without loss of relevant information as a result of the representation.

The use of semantic web techniques in education has also brought some improvements such as personalised content, recommendation of educational resources for students and the collection of data regarding students' interaction with educational environments [16]. This means that people can be matched to the educational offering available, which enables them to improve their knowledge and professional skills and help them in their decision making [11]. In addition, semantic web techniques have been combined with machine learning (ML) techniques. As a result, based on the available information, knowledge can be exploited and updated, and new relationships between the data can be inferred. The combination of an ontology-based recommendation system with ML techniques is emerging as an increasingly promising approach to improve the accuracy of recommendations.

Based on the above, and in light of the importance of developing professional skills, a hybrid recommendation system is proposed here. At its core is a semantic recommendation system that uses an ontology to model employment sectors and areas of knowledge to represent professional skills. The ontology is updated via events based on profiled data from professional records extracted from LinkedIn. ML is used to cluster entities in order to make predictions about new data. Subsequent stages make use of content filtering and heuristics to propose further lifelong learning courses for building and/or updating skills. The courses are selected from the course catalogue previously selected from the web, related to the user's job sectors, whose content increases the level of specialization and/or the degree to which the user's skills are current. This proposal is presented as a development of a hybrid recommendation system based on taxonomies and collaborative filtering [11] to test the use of ontology and ML in improving the performance and precision of RSs.

This article is structured as follows: Section 2 describes related studies; Section 3 outlines the proposed RS; Section 4 describes the evaluation of the proposed RS in detail; and Section 5 contains the conclusion and recommendations for future research.

2. Related Studies

The labour market is currently undergoing transformations driven by the changes brought about by the information society and the digital economy. Recently, the COVID-19 pandemic has had an impact on organisations, confronting them with new challenges and forcing them to adapt in order to survive [17]. Employability requirements have changed to address the new situations and challenges [1]. Employees not only need excellent technical skills, they also need sufficient skills to adapt to the changing requirements of their job. This means that they are expected to efficiently select knowledge from the vast amount of information available and effectively apply that knowledge [5].

As the amount of data available on the web grows, and users need to spend hours to find the information that meets their interests, there is a concurrent need to provide personalised and relevant recommendations. RSs have emerged to support this process. These are programmes that collect information that links users to items and use it to make relevant and meaningful suggestions based on the prediction of user interest, according to explicit or implicit preferences [18]. One of the areas where these systems can play an important role is in helping users achieve their career goals by generating personalised recommendations for job offers and skills development [19].

In the literature review, the study discussed by [20] proposed a system that supports users in designing their lifelong learning pathway according to professional expectations. The proposed system uses Natural Language Processing (NLP) techniques to process job offers automatically and extract the skills and knowledge required for each offer. This enables users to identify those that they need to develop, and ultimately recommends the courses that can be taken to acquire them and meet employment opportunities.

In order to recommend courses to students for their future careers, [21] proposed a taxonomy-based hybrid recommendation method. The system proposed three taxonomies: one for student profiles, one for courses and one for job offers. Tests showed that the use of taxonomies can improve recommendation results and help combat information overload and cold start problems. These authors proposed that further research could incorporate context into user information in order to make the system more comprehensive and intelligent. Likewise, [22] proposed a taxonomy for the representation of job descriptions and profiles of LinkedIn and Facebook users with information on education and work. The aim of the taxonomy proposal was to use it in an RS that applies heuristics, support vector machines and taxonomic similarity functions for job recommendations. These authors suggested that areas for future study could include metadata on users, and handle languages other than English.

The study presented by [11] proposed a hybrid RS for recommending lifelong learning courses aimed at improving and developing professional competences based on information on the employment sector and skills of LinkedIn users. A taxonomy code was used to classify skills by area of knowledge, taking into account the level of specialisation and degree of updating, which meant that more efficient comparisons could be made, and similarities could be calculated, in order to offer a better recommendation. The system tests showed a 20% increase in calculating similarities when using a taxonomy rather than a flat list, with the consequent improvement in the system recommendations. Areas of further research they proposed were improving the processes involved in retrieving contextualised information from users on LinkedIn and incorporating machine learning techniques.

The choice of data sources, the quality and the representation of data are of utmost importance for the good performance of RSs. According to [16], these systems basically need three components to provide recommendations: (1) context data (2) input data, namely, user information for the system to make a recommendation, and (3) an algorithm used in the recommendation process that operates jointly the context and input data to make recommendations for users. The basic idea of the whole system is not only the resulting recommendation algorithm, but mainly the creation of the entire process, from data acquisition through data processing to further use [23], which makes it possible to build user profiles including implicit and explicit information.

In this line of research, some studies can be found that seek to make recommendations more accurate by combining different filtering techniques with the use of ontologies in order to create a model that can describe the relationship between a domain area and the users that belong to it (Table 1).

Table 1. RSs that use ontology.

Ref.	Year	Objective	Techniques	Ontology	Assessment	Results
[24]	2015	Job recommendations	Ontology Similarity algorithms	Job vacancies Tasks	Not provided	Not provided
[25]	2015	Social media recommender systems for the development of social media skills	Ontology Filtering by content and context	Social media	Quality of results and integration of functions by users	3.62/5
[26]	2017	Educational resources recomendations	Ontology Sequential pattern mining (SPM)	Students' educational resources	Comparison of user satisfaction between the proposed system and other RSs, without ontology, and without SPM.	94%
[27]	2017	Recommendation of massive learning activities	OntologyKnowledge-based filtering	Students' Learning Activities	Not provided	Not provided
[28]	2017	Recommendation of learning materials	Ontology Canopy-K means Collaborative filtering	Educational Material	Comparison with other algorithm	Improved precision by 2.8

Table 1. *Cont.*

Ref.	Year	Objective	Techniques	Ontology	Assessment	Results
[29]	2018	Job recommendations	Ontology	Skills	Calculating similarity scores of Pointwise Mutual Information (PMI)	0.975
[30]	2018	Career path recommendation	Ontology	Skills	Calculation of information recall related parameters	Precision: 80.54% Recall: 86.44%
[31]	2019	Recommending the right degree and university for each person	Ontology K-means	Higher education institution. Student Work	Proposal	Proposal
[32]	2020	Recommendation of e-learning resources	Ontology Content-based filtering	Student Learning content	Proposal	Proposal

A review of the literature on the use of ontologies for RSs highlights the importance of information representation and management. The study by [25] indicated that the amount of information handled is a key factor for recommendations to be successful and proposed that future studies could do tests with real data in order to have a sufficient sample size.

Likewise, [24,26,27,30] pointed out that ontologies make it possible to infer information of interest to users, which can be used to enrich recommendations. Study [28] concluded that they take advantage of the similarity function to find additional information.

Studies [26–29] indicated that the use of the ontology improved the cold start and data sparsity problems that limit the performance of RSs. They also noted that personalising the learner profile through the use of ontology means that recommendations can be better tailored to the preferences of the target learners. In [27], the recommendations are shown in a pop-up window of the application, while the output of the system [30] is represented as a sequence graph based on the candidate skills. Study [32] concluded that knowledge representation is a vital step in building a knowledge-based recommendation for machine learning; therefore, it is essential to represent information in a machine-readable structure. Study [31] also suggested that the use of ontologies facilitates the analysis, reasoning, sharing and reuse of knowledge.

In relation to combining several techniques for recommendations, [27] concluded that the main benefit is to leverage the strength of each technique to overcome the limitations of the individual techniques on their own. Study [26] proposed that further research could include other intelligent tools and technologies from fields such as data mining and machine learning in the recommendation process to further improve the performance and precision of recommendations.

Based on this review and on the advances obtained in [11], it is proposed to use an ontology to enrich the existing taxonomies in order to model the knowledge related to the users' skills and work sectors to optimise performance and optimise the prediction of the recommendation system for lifelong learning courses. Using profiling data from professional records in social media, the skills used in different positions and professional practice (job sectors) were determined. The latter were then grouped using ML techniques to predict affinities between a new user profile, job sectors, and skills modelled on the ontology. Constant changes in the domain and context were managed via ontology updates and similarity clustering of positions using training data. For the recommendation process, a three-stage hybrid system is proposed: semantic filtering that makes use of the ontology that determines the skills to be updated and/or developed; a filtering by content for the initial prediction of courses; and heuristics to obtain the course offering to be recommended.

3. Materials and Methods

3.1. Materials

After analysing several SMSs, LinkedIn was selected because users share information about their jobs, careers and skills, and it is one of the most widely used SMSs in candidate selection processes [10]. LinkedIn user profiles were extracted from raw records. The terms found and the different contexts in which they appeared were analysed. These were initially used to build two ontologies, one to represent the areas of knowledge associated with skills, and the other to represent employment information. Both served as a basis for defining the heuristics for profiling the records.

For the user profiles, LinkedIn user records related to the area of software development were used to identify the data relevant to the objective of the research problem. The information extracted from LinkedIn for this purpose was: user ID, headline, location, brief educational background, skills, industry and interests.

In the context of the era of big data, the relationship between the data to be processed increases geometrically with the amount of data [33]. For improved processing power, the ontology instances were implemented as a knowledge graph. Neo4j, which is a NoSQL graph database manager, was used. Neo4j handles nodes and relationships and allows efficient operations on graphs. According to [33], the relationship is the most important element of the database in Neo4j, as it represents the interconnection between nodes, which makes it ideal for representing knowledge graphs. The different programmes that make up the system were programmed with Python 3. Cypher, the Neo4j query language, was used to query the graph.

The data obtained in [11] were used to evaluate the results of this proposal. This allows the behaviour and performance of the recommender system to be compared and assessed.

3.2. Methods

A hybrid filtering based on an ontology and ML was proposed for recommending lifelong learning courses for LinkedIn users. This was where the ontology was updated via events from user profiles, and ML was used to cluster different related entities to enrich the ontology. For this proposal, DBSCAN and K-Means clustering algorithms were used for evaluation purposes.

The objective of our proposal was to improve the performance and quality of the recommendations of [11] by moving from a system with stored data to an RS with a flexible architecture. This was based on a model that allows working with job sectors and user skills, which improves the recommendations for updating or developing skills by using taxonomic similarity and machine learning. The aim was to improve users' professional profile in a highly changing and competitive work environment.

For the recommendation process, a three-stage hybrid system was proposed: semantic filtering that makes use of the ontology that determines the skills to be updated and/or developed; a filtering by content for the initial prediction of courses; and heuristics to obtain the course offering to be recommended. The architecture of the proposed RS is presented in Figure 1.

The system is composed of two major phases: an off-line phase, in which the definitions are uploaded, and the models are built, and an on-line phase in which a recommendation of courses from a previously profiled offer for the improvement and/or development of professional skills is predicted on the basis of a user's record.
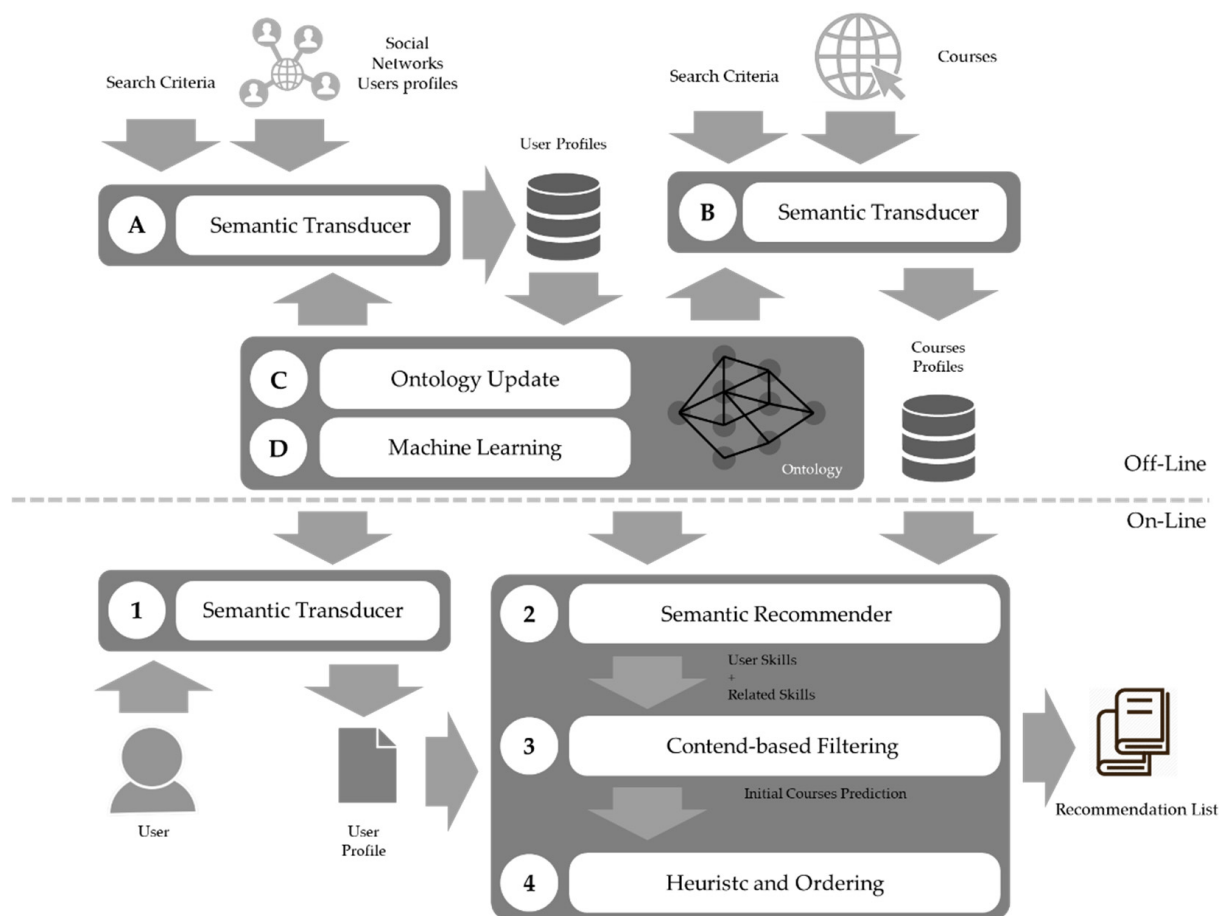
Figure 1. Recommendation System Flow.

3.2.1. Off-Line Phase

A key element in good RS performance is ensuring the richness of semantic data mining and avoiding the loss of information obtained from data recall. This phase aims at building data models to represent the domain of skills and job sectors by knowledge areas.

Ontology

An ontology was built to model the entities and relationships between job titles and knowledge through skills and the grouping of job titles by affinities. It was based on the taxonomies used in [11], where an area of knowledge and job positions are represented hierarchically and enriched with new relationships. Figure 2 shows the different entities of the ontology: knowledge areas, job sectors, skills, requirements and clusters.

Knowledge areas were represented hierarchically according to area, sub-area, speciality, subspeciality and specific knowledge. Job sectors (job titles or positions) were represented using a hierarchy consisting of sector, field, type, domain and specific job title. Each job sector has inherent skills for the position, which in themselves have an associated level of specialisation and degree of updating and belong to a specific area of knowledge. Job requirements had as properties a type, a description and a value in the case of being quantifiable. The clusters that represented groupings (under different criteria or algorithms) of related positions according to the similarity of the set of skills they make use of, required an associated set of skills (centroid) for prediction purposes.

For knowledge areas, the 'is equivalent to' relationship was introduced to model synonyms, and the 'is of interest' relationship was introduced to model relationships of interest that were complementary or similar to specific knowledge.
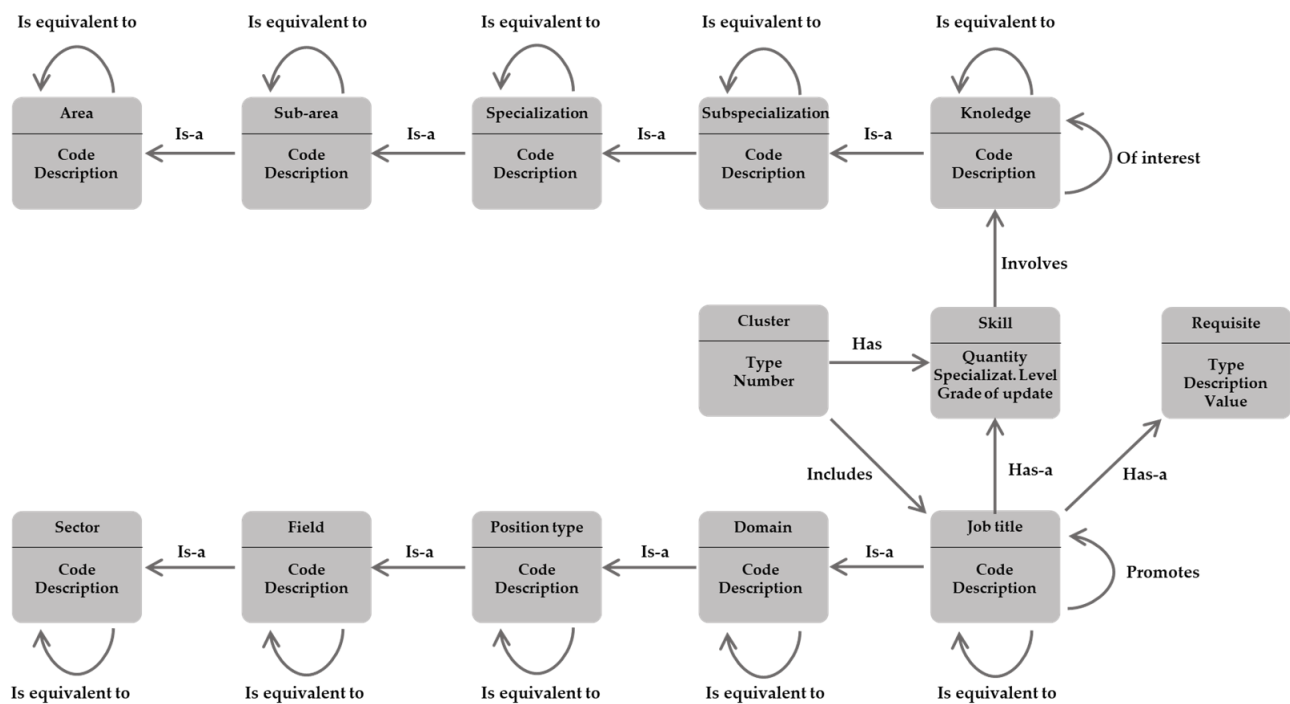
Figure 2. Ontology used in the RS.

For the job sectors, the entity requirements and the 'has' relationship were added to model the requirements associated with the positions and additionally the 'promotes' relationship was introduced to model the promotion relationships between positions in a job hierarchy.

For the purposes of the RS, the existing relationships between job titles and the skills possessed by the people who hold them (areas of knowledge, average level of specialisation and degree of updating) and frequency with which they appear, were modelled.

Finally, entities were modelled to represent the clusters obtained by applying ML techniques (including for K-means the representation of the centroids for the associated skills). Then, the user and course profiles were built using transducers (A and B) that make use of the ontology, which had previously been loaded as a knowledge graph from definition files. Based on the user profiles, the ontology was updated via events (C), i.e., the model was updated with the relationships between work sectors and areas of knowledge according to the skills of the user profiles. Finally, using machine learning (D), instances of entities in the ontology were clustered by using DB-SCAN and K-means algorithms.

3.2.2. On-Line Phase

This is the stage when the recommendation is made. This involves two processes involving a semantic transducer (1) that takes the user's record and generates their profile, and the hybrid recommendation system composed of three stages: a semantic recommender (2), content filtering (3) and a three-step heuristic (4), which refines the filtering process and generates the final recommendation for the user.

User Profiles

The user profile information consisted of the following elements: demographic information, job sectors and skill sets. Demographic information comprised a user ID, LinkedIn URL, place of residence, language skills and level of proficiency in each language and academic level. The job sectors resulted from a hierarchical code defined by the ontology, which was composed of: sector, field, job type, domain and job title. Finally, there were the set of skills in which each skill was represented by three arguments: a hierarchical

code defined by a taxonomy of knowledge areas, level of specialisation and degree of skill updating (Figure 3).
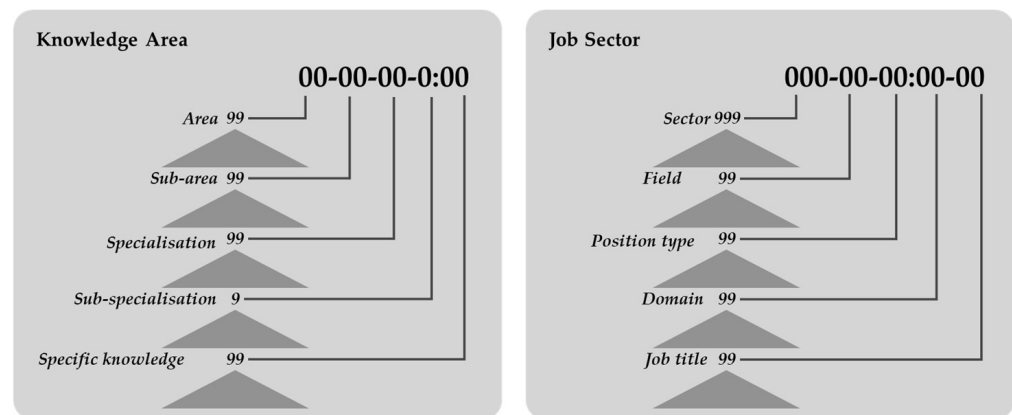

Figure 3. Hierarchical encodings of User Profiles.

Course Profiles

Course profiles consisted of demographic information and assessment of the course, skills to be acquired and skills required. The same coding used for the areas of knowledge in the user profile skills were employed for the areas of knowledge for the skills to be acquired and the prerequisites, which facilitated the filtering processes.

The demographic information and assessment of the course contained its name and description, URL on the WEB, locations, languages and dates available, as well as reviews of the institution and the course, if any.

The skills to be acquired in a course were a list of the skills expected to be obtained upon successful completion of the course, where each of them was represented by a hierarchical code defined by a taxonomy of knowledge areas and a level of specialisation. The skills required were related to a set of skills represented by a hierarchical code that was defined by a taxonomy of knowledge areas.

Similarity Functions

The similarity between user profiles and course profiles needed to be calculated for each of the recommendation algorithms. Likewise, machine learning algorithms used the distance between training elements for model building.

A path length was used to calculate the similarity of two coded elements using a taxonomy. In the case here, it was proposed to weight each of the levels, and a threshold was defined at which the elements could be said to be similar. It should be noted that given the similarity, the distance was defined as the difference at 1.0. For comparison and filtering purposes the following similarity functions were defined.

Taxonomic Similarity for Job Sectors

This compared two elements belonging to job sectors and was defined according to (1):

$$JS\_TS(js_i, js_j) = \sum_{l=1}^{D(T)} \left(K_l * JS\_Match(js_i, js_j, l) \right) \tag{1}$$

where:

js_i, js_j Job sectors.
$D(T)$: Depth of taxonomy.
K_l: Constant for relevance of level l in the taxonomy.
$JS\_Match(js_i, js_j, l)$: Level l function comparison, with wildcard support and $F : js \times js \times l \rightarrow [0, 1]$

By calculating $JS\_TS(js_i, js_j)$, the weighted length of the path sharing js_i and js_j in the taxonomywas verified for $l > 1$ that $JS\_Match(js_i, js_j, l - 1) > 0$, otherwise $JS\_Match(js_i, js_j, l) = 0$.

For the purpose of experimentation, it was proposed that the following initial values be used for level l relevance constants, $K_l = \frac{1}{D(T)}$, where it should be verified that $\sum\limits_{l=1}^{D(T)} K_l = 1.0$.

Taxonomic Similarity for Areas of Knowledge

This compares two elements belonging to areas of knowledge and was defined according to (2):

$$KA\_TS(s_i, s_j) = \sum_{l=1}^{D(T)} \left(K_l * KA\_Match(KA(s_i), KA(s_j), l) \right) \tag{2}$$

where:

s_i, s_j: Skills for users *i* and *j*.
$KA(s)$: Area of knowledge for a particular skill.
$D(T)$: Depth of taxonomy.
K_l: Constant for relevance of level l in the taxonomy, where it should be verified that $\sum\limits_{l=1}^{D(T)} K_l = 1.0$.

$KA\_Match(ka_i, ka_j, l)$: Level l function comparison, with wildcard support and $F : ka \times ka \times l \rightarrow [0, 1]$.

By calculating $KA\_TS(s_i, s_j)$ the weighted length of the path sharing $KA(s_i)$ and $KA(s_j)$ in the taxonomy, it was verified for $l > 1$ that $KA\_Match(KA(s_i), KA(s_j), l - 1) > 0$, otherwise $KA\_Match(KA(s_i), KA(s_j), l) = 0$.

For the purpose of experimentation, it was proposed to use the following initial value for the level l relevance constants: $K_l = \frac{1}{D(T)}$.

Similarity of User Skills

The cosine similarity function or Pearson's similarity correlation function was used to compare the skill sets of different users. To do so, skill vectors must be calculated over the same *n*-dimensional space. In order to handle two magnitudes, it was proposed to use assessment of skill development.

Assessment of Skill Development

This allowed the same skill in different profiles to be compared, based on two attributes, level of specialisation and degree of skill updating, for which the Euclidean distance given by (3) was used:

$$KV(s) = \sqrt{SL(s)^2 + UG(s)^2} \tag{3}$$

where:

$SL(s)$: Specialisation level of a skill.
$UG(s)$: Degree of skill updating.

It is common for users to have different skill sets. In order to calculate similarity, it is useful to represent these as vectors in the same *n*-dimensional space; however, the cardinality of the skill set can be high and representing all dimensions would result in many zero entries. In addition, as the areas of knowledge are represented by a hierarchical code in a taxonomy, comparisons at different levels of the taxonomy can be useful.

Therefore, it is useful to construct the vectors for knowledge ratings, taking into account the taxonomic similarity function for the areas of knowledge, which allows the

comparison of related skills with a cut-off threshold of less than or equal to 1.0, allowing all related skills to be grouped under the same entry. For this purpose, the maximum value was used as a function summary for the assessment of skill development. For the construction of the vectors, it was verified that:

$$S = S_i \cup S_j$$
$$[s] = \{ s' \in S, \ s' \sim s \}$$
$$[s] = \{ s' \in S, \ KA\_TS(s,s') \geq \theta, \}$$
$$\tilde{S} = \{ [s], \ s \in S \}$$
$$V, U = uv\_Cal(S_i, S_j, \theta)$$

where:

S_i: Skill set *i.*
S_j: Skill set *j.*
θ: Cut-off threshold for the taxonomic similarity function.
$V_k = max(KV(s_l)), \ s_l \in S_i, \ s_l \in \tilde{s}_k, \ \tilde{s}_k \in \tilde{S}.$
$U_k = max(KV(s_l)), \ s_l \in S_j, \ s_l \in \tilde{s}_k, \ \tilde{s}_k \in \tilde{S}.$

The vectors V and U were constructed from the quotient set obtained by defining the taxonomic similarity function with the cut off threshold θ over the set $S_i \cup S_j$ as the equivalence relationship, where each element with an index $k \ for \ V$ and U was the maximum value for the assessment of skills for S_i and S_j, where it was verified that the skills belonged to the k- equivalence class of the quotient group.

Cosine Similarity Function for User Skills

The cosine similarity function based on the vectors for skill development ratings was calculated according to 4:

$$Cos\_Sim(v,u) = \frac{\sum_{i=1}^{n}(v_i * u_i)}{\sqrt{\sum_{i=1}^{n} v_i^2} + \sqrt{\sum_{i=1}^{n} u_i^2}} \tag{4}$$

where:

v, u: Vectors for skill development ratings.
$n = \mid v \mid = \mid u \mid.$

Pearson's Similarity Function for User Skills

Pearson's similarity function was defined based on the vectors of skill development ratings, as (5):

$$Pearson\_Sim(v,u) = \frac{\sum_{i=1}^{n}((v_i - \overline{v}) * (u_i - \overline{u}))}{\sqrt{\sum_{i=1}^{n}(v_i - \overline{v})^2} + \sqrt{\sum_{i=1}^{n}(u_i - \overline{u})^2}} \tag{5}$$

where:

v, u: Vectors for skill development ratings.
$n = \mid v \mid = \mid u \mid.$

Recommendation Process

The recommendation process consisted of three stages, which are described below.

Determining Related Job Sectors (2)

In this first stage of the recommender, the user's profile was used to determine the user's possible related job sectors (RJSs) from their user's job sectors (UJSs). This was done by using the knowledge graph and predicting the clusters to which they belong. For the clustering prediction for a new piece of data when using the DBSCAN algorithm, a core point that was less than epsilon distance was located and linked to the cluster it belongs to. For K-means, the centroid with the smallest distance was located and linked to that cluster.

In both cases 1 minus the similarity function, was used as the distance (6), using either (4) or (5) for the calculation of similarity:

$$d = 1 - Sim(S_u, S_c) \tag{6}$$

It was then determined that related job sectors (RJSs) were prior to user job sectors (UJS) in an orderly relationship imposed by a hierarchy and were removed from the related job sectors. Subsequently, the related skills were calculated (using their current and related job sectors), which were the skills most frequently associated with related job sectors that add value to the user's skills (USs), i.e., the skill does not exist in the set of skills or, otherwise, it has a higher level of specialisation or degree of updating. Finally, those skills that were of interest in the knowledge graph were added to the related skills.

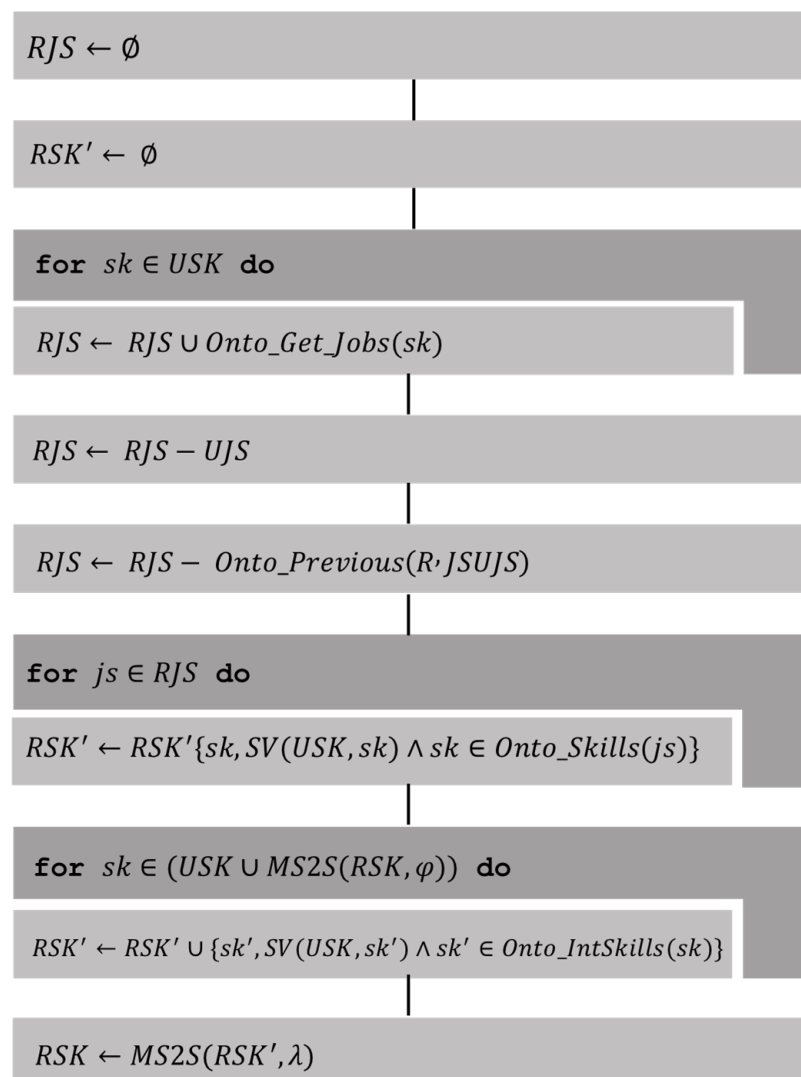The following algorithm (Figure 4) was used to determine the related user skills.

Figure 4. Related Job Sectors algorithm.

Alternatively, to determine related skills using ML clustering, the following algorithm (Figure 5) was used.
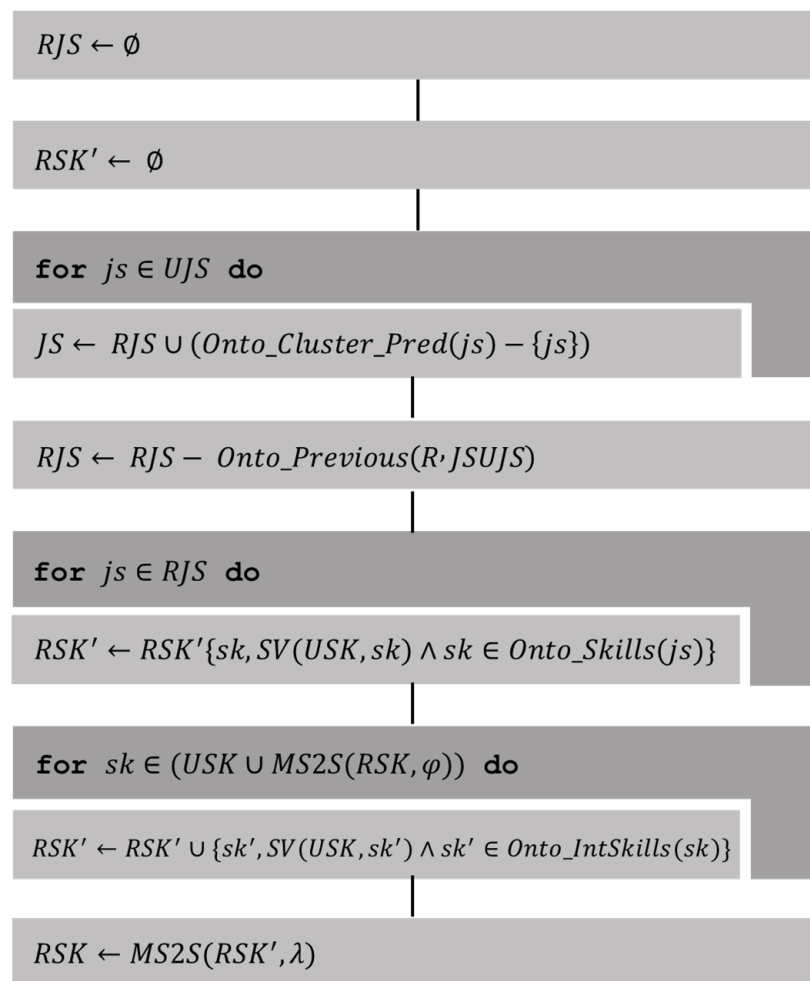
$$RJS \leftarrow \emptyset$$

$$RSK' \leftarrow \emptyset$$

for $js \in UJS$ **do**

$JS \leftarrow RJS \cup (Onto\_Cluster\_Pred(js) - \{js\})$

$RJS \leftarrow RJS - Onto\_Previous(R^{,}JSUJS)$

for $js \in RJS$ **do**

$RSK' \leftarrow RSK'\{sk, SV(USK, sk) \wedge sk \in Onto\_Skills(js)\}$

for $sk \in (USK \cup MS2S(RSK, \varphi))$ **do**

$RSK' \leftarrow RSK' \cup \{sk', SV(USK, sk') \wedge sk' \in Onto\_IntSkills(sk)\}$

$RSK \leftarrow MS2S(RSK', \lambda)$

Figure 5. Related skills using ML clustering algorithm.

where:

$Onto\_Get\_Jobs(sk)$, job sectors where users have the skill sk recorded through events in the ontology.

$Onto\_Cluster\_Pred(js)$, predicts the cluster for a job sector js according to clustering by ML in the ontology.

$Onto\_Previous(RJS, UJS)$, the previous sectors in RJS to UJS according to the orderly relationship given by the hierarchy in the ontology.

$Onto\_Skills(JS_i)$ are the JS_i skills represented in the ontology, which are the product of an event-driven update of user profiles.

$Onto\_IntSkills(sk)$ are the skills of interest associated with the sk sk skills represented in the ontology.

$SV(SK, sk)$, function that verifies that the skill sk is valuable for the skills SK.

$MS2S(MS, c)$ constructs a set with the most frequent elements c c in the multiset MS.

RSK', Multiset of related skills.

Initial Course Prediction (3)

Using content filtering, courses were selected from the course catalogue, previously extracted from the web related to the user's area of expertise, according to whether they provide a higher level of specialisation and/or the degree of updating of the individual's own skills or help to develop the user's related skills.

The following algorithm (Figure 6) was used to determine the initial course prediction for the user.
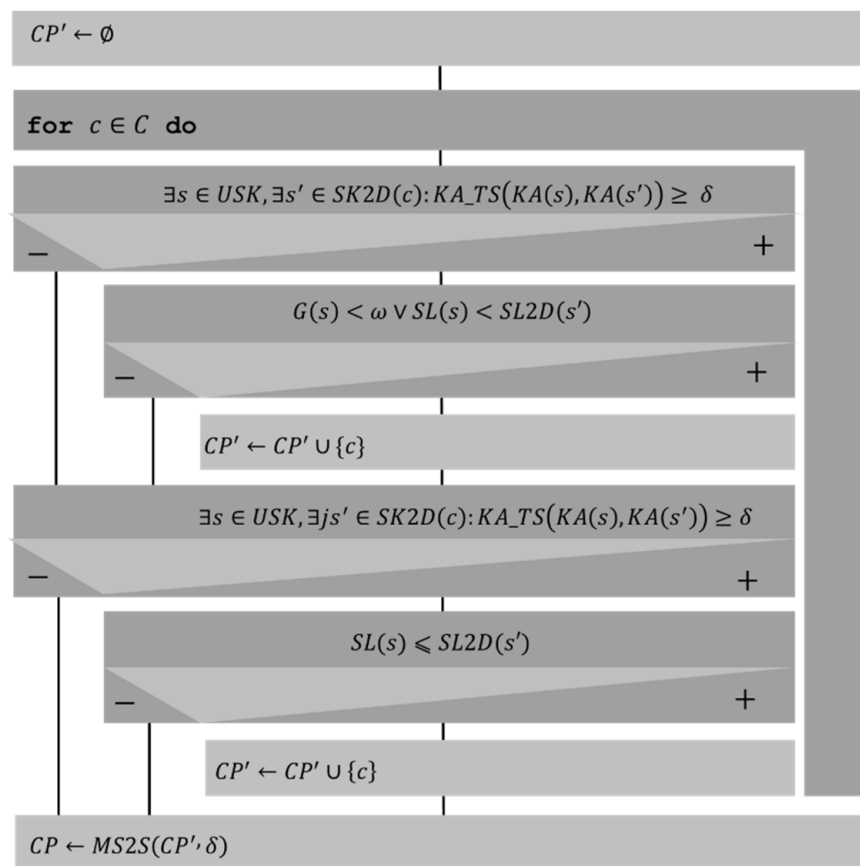
Figure 6. Initial course prediction algorithm.

where:

$SK2D(c)$, skills to be developed in the course c.

$KA(s)$, skill knowledge area s.

$SL(s)$, skill specialisation level s s.

$UG(s)$, degree of updating of skills s.

$KA\_TS(sk_i, sk_j)$, taxonomic similarity between areas of knowledge sk_i, sk_j.

$MS2S(MS, c)$ constructs a set with the most frequent elements c in the multiset MS.

CP', multiset of course prediction for the user.

Filtering Heuristics and Course Ordering (4)

Finally, the course prediction of the previous phase was complemented and filtered using heuristics. Given the user profile and the course catalogue, it was verified that.

1. Users possessed the necessary skills to take the courses; otherwise, courses were selected to develop those skills, applying demographic restrictions.
2. User demographic restrictions were applied to the initial course prediction, and to the result of the previous phase.
3. In the prediction of courses, courses where the skills to be developed were the same or a subset of another were deleted, and the best-rated prediction was kept.

The following algorithm (Figure 7) was used to determine the course recommendation for the user.
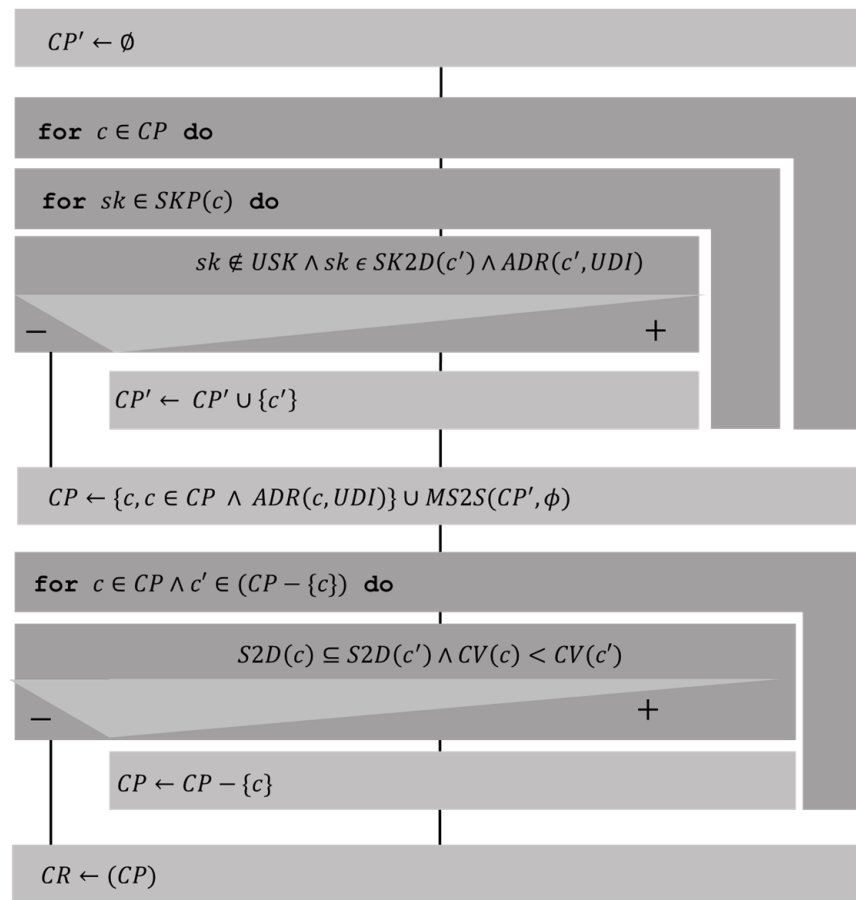
$$CP' \leftarrow \emptyset$$

for $c \in CP$ **do**

for $sk \in SKP(c)$ **do**

$$sk \notin USK \wedge sk \in SK2D(c') \wedge ADR(c', UDI)$$

$-$ $+$

$$CP' \leftarrow CP' \cup \{c'\}$$

$$CP \leftarrow \{c, c \in CP \wedge ADR(c, UDI)\} \cup MS2S(CP', \phi)$$

for $c \in CP \wedge c' \in (CP - \{c\})$ **do**

$$S2D(c) \subseteq S2D(c') \wedge CV(c) < CV(c')$$

$-$ $+$

$$CP \leftarrow CP - \{c\}$$

$$CR \leftarrow (CP)$$

Figure 7. Final courses prediction algorithm.

where:

$SKP(c)$, prerequisite skills for the course c.

$SK2D(c)$, skills to be developed in the course c.

$ADR(c, di)$, applies demographic restrictions di to the course c.

$CV(c)$, Course *ratings c*.

$MS2S(MS, c)$, constructs a set with the most frequent elements c in the multiset MS.

CP', Mutiset course prediction for the user.

Measurements for the Evaluation of RSs

According to [34], while the precision of recommendations is an important factor in measuring the performance of RSs, it is not sufficient on its own. Users may also be interested in preserving their privacy, in the response time of the system, in discovering new items, as well as the order in which they are presented, and other properties related to system interaction. The evaluation system must be designed carefully so that the metrics measured really reflect the effectiveness of the system from the user perspective. Measurements such as novelty, trust, coverage and serendipity are important for user experience, as these have an important short and long-term impact on conversion rates [12].

In order to evaluate the effectiveness of the RS proposed here, a series of measurements for evaluating recommender systems were used, which are defined in Table 2 below.

Table 2. Recommendation System evaluation measurements.

Measurement	Description	Calculation for the User	Calculation for System
MAE	Mean absolute error of the given recommendation vs. the recommendation expected by the user	$MAE_i = \frac{|(CE_i \cup CR_i) - (CE_i \cap CR_i)|}{|CE_i \cup CR_i|}$	$MAE = \frac{\sum_{i=1}^{n}|(CE_i \cup CR_i) - (CE_i \cap CR_i)|}{\sum_{i=1}^{n}|CE_i \cup CR_i|}$
RMSE	Root mean square error	$RMSE_i = \sqrt{\frac{|(CE_i \cup CR_i) - (CE_i \cap CR_i)|^2}{|CE_i \cup CR_i|}}$	$RMSE = \sqrt{\frac{\sum_{i=1}^{n}|(CE_i \cup CR_i) - (CE_i \cap CR_i)|^2}{\sum_{i=1}^{n}|CE_i \cup CR_i|}}$
Coverage	Users to whom the system has made a recommendation, CR_i set of recommendations from $u_i \in U$		$C(S) = \frac{|\{u_i \in U, CR_i \ \varnothing\}|}{|U|}$
Precision:	The fraction of the recommendation that is relevant to the user.	$P(u_i) = \frac{|\{c \in CR_i, relevante(c,u_i)\}|}{|CR_i|}$	$P(S) = \frac{1}{|U|} \sum_{u_i \in U} P(u_i)$
Recall:	Calculates the ratio between recommendation and user preference, where CE_i is the preference group of $u_i \in U$	$R(u_i) = \frac{|CE_i - CR_i|}{|CE_i|}$	$R(S) = \frac{1}{|U|} \sum_{u_i \in U} R(u_i)$
Novelty	The portion of recommendations made to the user that the user is not familiar with or has not seen before	$N(u_i) = \frac{|CR_i - CE_i|}{|CR_i|}$	$N(S) = \frac{1}{|U|} \sum_{u_i \in U} N(u_i)$
Serendipity	The fraction of the recommendation that is unexpected and valuable to the user.	$S(u_i) = \frac{|\{c \in CR_i, serendipity(c,u_i)\}|}{|CR_i|}$ Where: $\{c \in CR_i, serendipity(c,u_i)\} \subseteq (CR_i - CE_i)$	$S(S) = \frac{1}{|U|} \sum_{u_i \in U} S(u_i)$

4. Results

The initial loading and updating of the knowledge graph was performed based on the ontology definitions and user data. For each of the user profiles, the skills associated with the job sectors were updated through events. In order to determine the clusters, the programmes implementing the machine learning algorithms were then run: K-means and DBSCAN. Figure 8 shows, in part, the entities and relationships associated with the iOS native mobile application developer job sector.

Once the data model was loaded, the RS was tested. After inputting a user profile and then applying the different filtering stages, the course prediction was performed. For evaluation purposes, the application was run using different configurations, which are briefly described below.

1. Content filtering using user's own and associated skills; no related skills were determined.
2. Collaborative filtering using only user's own job sectors; semantically related skills were determined.
3. Semantic filtering using rules to determine related job sectors and related skills.

4. Semantic filtering using 75% coverage of user skills to determine related job sectors, i.e., those job sectors that covered 75% of the user's skills were selected and used to determine related skills.
5. Semantic filtering using 50% coverage of user skills to determine related job sectors, i.e., those job sectors that covered 50% of the user's skills were selected and used to determine related skills.
6. Semantic filtering using DBSCAN clustering to determine related job sectors and semantic rules to establish related skills.
7. Semantic filtering using K-Means clustering to determine related job sectors and semantic rules to establish related skills.
8. Semantic filtering using DBSCAN and K-Means clustering to determine related job sectors and semantic rules for related skills using the predicted job sectors with both clusters, and related skills were determined based on both clusters.

Figure 8. Extract from the knowledge graph.

For each of the test cases, a filtering of the related work sectors was carried out to discard those prior to the user's current job sectors, according to the orderly relationship imposed by a hierarchy; likewise, the related skills were enriched with the skills of interest described in the ontology. Cosine similarity was used and an ontological similarity function with a threshold cut-off at 0.8 was used in the construction of the U and V vectors.

Figure 9 shows the output of the RS using the K-means clustering prediction to determine related job sectors corresponding to configuration 7.

Three tests were performed to evaluate the performance of the system for each of the configurations, which corresponded to three different sets of data. In each run the system recommendations and the preferences of each user were used to estimate user measurements; measurements for each user led to the estimation of the system measurements.

```
================================================================
                    Hybrid Recommemnder System
----------------------------------------------------------------
                       Clustering / K-Means
================================================================
                    Semantic Recommender Output
----------------------------------------------------------------
  User Job Sectors:
       001-03-07:03-02 : iOS Nativo
  User skills:
       ('20-10-01-*:01', 4, 2) : Xcode
       ('10-06-01-5:03', 4, 2) : ObjetiveC
       ('10-07-09-2:01', 3, 2) : MVC
       ('20-09-02-*:00', 3, 3) : Frontend
  Previous Job Sectors:
  Related Job Sectors:
       001-03-08:00-04 : Code Reviewer
       001-03-07:03-02 : Native iOS
       001-03-07:03-00 : Mobile Applications
  Related Skills:
       ('10-06-01-9:02', 3, 3) : Swift
       ('10-07-09-2:03', 3, 2) : MVVM
       ('10-07-09-2:04', 3, 2) : VIPER
       ('10-07-05-*:01', 3, 2) : Mantis
       ('10-06-01-5:05', 3, 3) : Kotlin
       ('10-07-09-2:02', 3, 3) : MVP
       ('20-10-02-*:01', 3, 2) : AndroidStudio
       ('10-06-01-5:03', 4, 3) : ObjetiveC
  Associated Skills:
       ('10-07-07-*:00', 0, 0) : Version management
       ('20-09-02-1:00', 0, 0) : Progressive
================================================================
           Content Filtering Output (Initial Prediction)
----------------------------------------------------------------
   7  Developed in Android with Kotlin
   2  Developed in iOS with Swift
   3  Basic programming with Swift
   4  Introduction to programming with Swift - Bilbao
   5  Introduction to programming with Swift - Barcelona
   6  Advance development with Android
   8  Basics of programming with Kotlin
   9  Introduction to programming with Kotlin -  Bilbao
  10  Introduction to programming with Kotlin - Barcelona
  15  Version management with Gif
================================================================
             Heuristics output (course recommendation))
----------------------------------------------------------------
   7 Developed in Android with Kotlin
   2 Developed in iOS with Swift
   3 Basic programming with Swift
   6 Advance development with Android
   8 Basics of programming with Kotlin
  15 Basics of programming with Kotlin
================================================================
```

Figure 9. Recommendation system output.

The first test was carried out with training data, which corresponded to 70% of the data. The results are shown in Table 3. The second test corresponded to 100% of the data (Table 4). And finally, the system was run using the 30% of the data that was not used for training (Table 5).

Table 3. RS performance measurements with training data.

Setting	MAE	RMSE	Coverage	Precision	Recall	Novelty	Serendipity
1	0.58	9.13	0.64	1.00	0.36	0.14	0.09
2	0.51	8.22	0.73	0.97	0.41	0.21	0.07
3	0.25	4.44	0.91	0.83	0.82	0.46	0.07
4	0.22	3.65	0.82	0.93	0.71	0.45	0.07
5	0.38	6.62	0.91	0.80	0.74	0.39	0.07
6	0.18	2.92	0.91	0.91	0.80	0.52	0.07
7	0.49	9.25	0.91	0.67	0.63	0.31	0.04
8	0.30	5.24	0.95	0.80	0.80	0.48	0.04
9 [1]	-	-	0.91	0.91	0.70	0.52	0.06

[1] Baseline-hybrid recommender data stored using taxonomies [11].

Table 4. RS performance measurements with total data.

Setting	MAE	RMSE	Coverage	Precision	Recall	Novelty	Serendipity
1	0.59	11.05	0.61	0.99	0.37	0.11	0.06
2	0.50	9.49	0.71	0.97	0.45	0.17	0.05
3	0.25	5.16	0.94	0.82	0.87	0.39	0.06
4	0.23	4.49	0.81	0.94	0.73	0.39	0.05
5	0.34	7.03	0.87	0.84	0.75	0.34	0.06
6	0.27	5.63	0.94	0.84	0.86	0.45	0.05
7	0.51	11.72	0.94	0.62	0.65	0.25	0.04
8	0.35	7.64	0.97	0.75	0.86	0.42	0.03

Table 5. RS performance measurements with test data (not used in the training).

Setting	MAE	RMSE	Coverage	Precision	Recall	Novelty	Serendipity
1	0.62	6.24	0.56	0.96	0.41	0.04	0.00
2	0.47	4.75	0.67	0.96	0.56	0.07	0.00
3	0.23	2.64	1.00	0.80	1.00	0.24	0.04
4	0.26	2.63	0.78	0.96	0.78	0.26	0.00
5	0.25	2.60	0.78	0.94	0.78	0.21	0.02
6	0.43	5.42	1.00	0.67	1.00	0.28	0.00
7	0.56	7.24	1.00	0.49	0.69	0.12	0.06
8	0.45	5.79	1.00	0.62	1.00	0.28	0.00

According to [12], although there is a natural trade-off between precision and recall, an increase in recall does not always lead to a reduction in precision. One way to create a single measurement that summarises both precision and recall is the harmonic mean between precision and recall (F1), which is calculated with (7):

$$F = \frac{2 * Precision * Recall}{Precision + Recall} \tag{7}$$

In order to evaluate our system, the harmonic mean was calculated for the three runs (Table 6).

Taking into account mainly the measurements of precision, recall and errors, in general, on analysis of Table 3 (performance with training data) and Table 6 (harmonic mean comparison), it was observed that in relation to the baseline performance (hybrid recommender system with stored data based on taxonomy [11]), better performance was observed for the DBSCAN-based configuration (setting 6) and for the configuration based on semantic rules (setting 3), which suggests a good modeling of the data by the ontology and clustering performed by DBSCAN. K-means clustering showed poorer performance, which could be explained by the nature of the domain and the distribution of the positions in space. Better clustering (related positions) was generally found when using DBSCAN

than K-means in the graph analysis. Improvements in recall and serendipity measurements could be associated with use in the ontology and the estimation of associated skills or skills of interest and filtering of positions prior to the current one in the order relationship imposed by a hierarchy.

Table 6. Calculation of the harmonic mean (F) for RS runs.

Setting	Test 1	Test 2	Test 3
1	0.53	0.54	0.57
2	0.58	0.61	0.71
3	0.82	0.84	0.89
4	0.81	0.82	0.86
5	0.77	0.79	0.85
6	0.85	0.85	0.8
7	0.65	0.63	0.57
8	0.8	0.8	0.77
9	0.79	-	-

Similar behaviour was observed when analysing Table 4, corresponding to the test with the total dataset, and Table 6, which showed slightly better performance in DBSCAN configuration (setting 6) over the configuration based on semantic rules (setting 3).

From an analysis of Table 5 (behaviour with test data) and Table 3, a decrease in precision and an improvement in recall can be seen for the DBSCAN (setting 6) and the semantic rule-based (setting 3) configurations in relation to the previous scenarios, i.e., the size of the recommendations increased in such a way that it covered relevant recommendations but increased differences. However, looking at Table 6 (comparison of harmonic means), an improvement in the semantic rule-based configuration (setting 3) and a decrease in the DBSCAN-based configuration (setting 6), were found, although both were always above the baseline case (setting 9). This can be explained by the immunity of the ontology to the absence of data.

Finally, a number of remarks can be made regarding the comparison of the capabilities and performance of the proposed system with similar works previously reviewed.

The study proposed by [24] used the tasks performed by individuals to recommend jobs through a ranking process based on ratings. Our proposal makes use of the positions held by the individual and their skills, and on the basis of the skills that are most used by those who hold these positions, related positions are provided and the skills to improve or develop are predicted and recommended, together with the relevant courses. In addition, the system has the capability of recommending jobs using a filter on the percentage coverage of the user's skills from the skills of the people in that position. It is possible to establish the skills to be improved or developed and to obtain course recommendations for this purpose.

In [26] an ontology was used to determine skills which calculated cosine similarity to determine similar students, and used the ratings of educational resources to make predictions. Our system does not use course ratings by individuals, since accessing this information is problematic [11]; instead, it recommends courses that improve or develop skills according to the most frequent skills found among the people who currently hold these positions. In terms of system performance against the MAE benchmark, both recommender systems behaved similarly in comparison with the DBSCAN configuration (setting 6).

The study proposed by [29] determined the similarity between skills for job recommendation. However, our proposal not only checks the similarity between skills, using taxonomic similarity, but also has the ability to determine the similarity and distance between skill sets. This makes it possible to identify related jobs or is used by machine learning algorithms to cluster similar positions, and to determine the skills to improve or develop, which offers good predicting computational performance.

Similar results were found for the measures of precision, recall and F (harmonic mean), even with alternate recommender systems between the study proposed by [30] and ours. Additionally, [30] proposed that future research could use the skills graph to recommend professional growth options for a user. Our proposal does this by using an alternative scheme, as explained above; it establishes the skills to improve or develop based on related positions, which are determined by similarity or clustering, making use of similarity with the most common skills found in the positions according to the training data set.

It was observed that most of the proposals studied that used ontology, employed OWL and related query languages. We used a knowledge graph using Neo4j, which can be advantageous when working with large datasets, in particular if they come from social media. It also allows graph algorithms to be used to identify properties of interest in the graph (centrality, density, radius, etc.).

In addition to the recommendations and information obtained from the system run, one of the advantages of using a knowledge graph is that associated queries can be made to obtain information of interest related to the problem. For example, we can check which is the most demanded skill for a particular area. Figure 10 partially shows the knowledge graph for the area of Android mobile application development, and the highest density can be observed in the Android Studio skill.
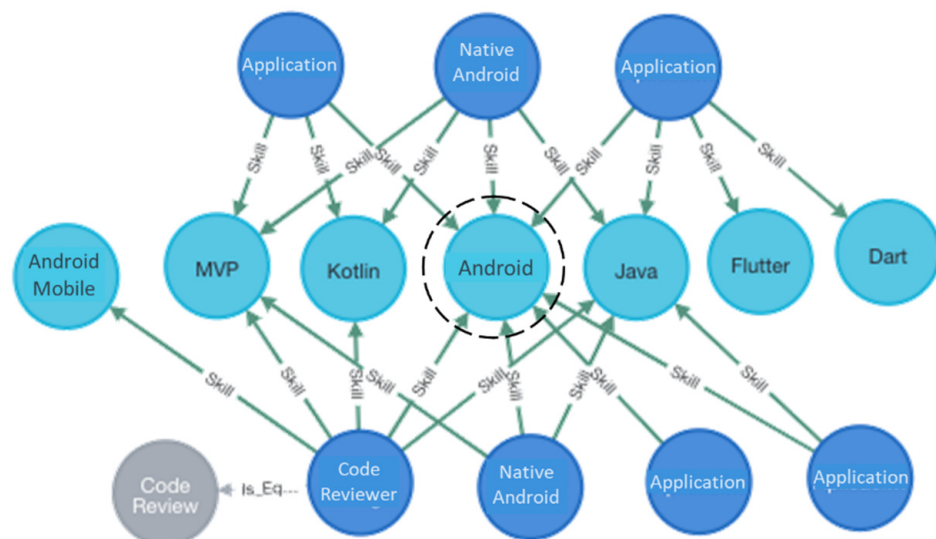

Figure 10. (Partial) knowledge graph for skills associated with Android mobile application development.

5. Discussion and Conclusions

In this paper we present a proposal that combines ontology with ML techniques, aimed at grouping related jobs according to the similarity of the skills they use, to make lifelong learning recommendations to professionals based on their skills and job sectors, current and related ones.

The system architecture made it possible to use different configurations and to evaluate each of the functionalities separately and, if necessary, several of them simultaneously, thanks to the use of accumulated frequency. The configurations can be parameterised, which allows a wide range of combinations to be evaluated.

The objective of the study proposed by [25] was the development of social media skills based on whether users had or lacked these skills. Our proposal not only works with the skills obtained by a user in their academic background or in the course of their job, but also considers their current position and related ones. On this basis, skills could be identified to be developed according to labour market trends, drawing on the skills that are most frequently found.

Although [27] did not show performance measurements for comparison purposes, it explained the advantage of having a hybrid recommendation system, particularly in

cold start if it makes use of a semantic stage. This advantage is evident in our proposal, with new profiles that do not have defined positions or that do not fall into any cluster, where the semantic filter was able to make a recommendation beyond that associated with a user's existing skills.

Unlike the study by [31], which proposed the use of a hybrid recommender using ontologies and machine learning techniques working with structured data, our proposal has the advantage of working with semistructured and unstructured data; while it outlines the system architecture in very general terms, it does not show performance information for comparison.

The results showed an overall improvement in the measurements of the RS proposed by [11], with the additional performance advantage of working with one model compared to the cost of processing all user data. Additionally, it is worth noting that given a job sector, the queries associated with predicting clustering were much simpler and less computationally expensive than determining related job sectors using semantic rules and knowledge graph inference.

Table 7 shows a comparison of the results obtained in our proposal with related work.

Table 7. Comparison of research results.

	[25]	[27]	[31]	[11]	Author's Proposal
Type	Hybrid RS	Hybrid RS	Hybrid RS	Hybrid RS	Hybrid RS
Items to be recommended	MOOC	MOOC	University courses	Lifelong learning	Lifelong learning
Use of social network	Yes	-	No	Yes	Yes
Semantic	Yes	Yes	Yes	No	Yes
Ontology	Yes	Yes	Yes	No	Yes
Machine learning	No	No	Yes	No	Yes
Types of data	-	-	Structured	Semi-structured/ Unstructured	Semi-structured/ Unstructured
Future scenarios	No	No	Yes	Yes	Yes
Harmonic mean	-	-	-	0.79	0.85 ML/DB-SCAN
Novelty/Serendipity	-	-	-	0.52/0.06	0.52/0.07 ML/DB-SCAN

Achievements included the following.

It has been shown that it is possible to move from a stored data philosophy to a semantic model representing the work domain, with ensuing computational benefits, where the training data are able to represent reality reliably, which can be evidenced by an improvement in the quality of the recommendations with respect to the results obtained in [11], especially in the metrics of novelty and serendipity. This is possible since the ontology represents associations between areas of knowledge (relationships of the type: "may be of interest", "complements" or "expands") as well as skills to be developed associated with related positions that are determined by labour market trends.

An RS was developed that can make use of a semantic model using ML techniques and modular construction, which allows multiple techniques and algorithms to be integrated into the recommendation, relying on cumulative frequencies.

Although this was not the focus of the proposal, skill coverage techniques make itpossible for the system to be used to recommend jobs, given a user profile, or to search for candidates or a job sector profile. They can also indicate what the labour market trends are, and the most demanded skills, through the analysis of the knowledge graph.

The use of semantic rules improves recommendations by improving refining of the filtering of positions (according to the hierarchy) and therefore of skills, as well as considering related skills. This is an improvement over [11], which suggested skills development courses for lower level positions due to it not consider job ranking or the option of user promotion.

With respect to the weaknesses found in our proposal, we highlight the fact that the ontology is not automatically updated. Future work may use ontologies from third parts that are updated frequently for job sectors and areas of knowledge.

Another aspect to improve is the user profile, incorporating learning style and incorporating the learner's preferences in terms of language or modality. For this purpose, future works could include information from other social networks to complement user profiles. The use of LinkedIn skills endorsements to improve recommendations could be considered.

Further research could involve working with several related domains (software development, technical support, consultancy, training and academia), in order to evaluate the behaviour of the system for multiple domains, and updating the ontology with new instances associated to new domains.

The differences in performance when using several clustering algorithms suggests considering other algorithms and evaluating the use of soft versus hard clustering, as well as reviewing the weighted or conditional use of multiple techniques and algorithms in the recommendation construction strategy, according to the nature of the user.

Reviewing the weighted or conditional use of multiple techniques and algorithms in the recommendation construction strategy could be evaluated according to the nature of the user.

The use of algorithmic techniques on the knowledge graph can be evaluated to improve recommendations based on the most in-demand positions and skills, and build pathways describing sequences to achieve skills.

In conclusion, we can highlight the following:

Based on the weaknesses and opportunities resulting from the evaluation of the recommender system based on stored data [11], it was proposed a hybrid recommender system based on semantics be developed, for which an ontology was built and updated using machine learning. The results obtained showed that the modeling represented reality quite well, that the use of an ontology (instead of a taxonomy and stored data) allowed good modeling of reality which not only resulted in a better quality of recommendations, but also in a simpler system with fewer steps, and which translated into a better performance in relation to the previous system based on stored data [11]. In the absence of training data, the use of ontology could generate acceptable recommendations; an expected result in the use of semantic recommenders.

The results show that the best performances were obtained when using semantic rules to determine related job sectors, and when making use of the positions in the predicted cluster within the set obtained by machine learning. In particular, a better performance of the density-based hard-clustering algorithm (DBSCAN) over the centroid-based one (K-Means) was observed. The differences in behaviour suggest considering other clustering algorithms and to consider evaluating soft-clustering algorithms in the future.

Making use of ML clustering is much simpler, computationally speaking, than determining related occupational sectors using semantic rules.

Finally, the architecture of the system, as well as the use of neo4J and Python, provided the flexibility and performance needed to evaluate multiple configurations and machine learning algorithms, providing a powerful platform for evaluating new algorithms and increasing the amount of training data in the future.

References

1. Raitskaya, L.; Tikhonova, E. Skills and Competencies in Higher Education and Beyond. *J. Lang. Educ.* **2019**, *5*, 4–8. [CrossRef]
2. Kenayathulla, H.B.; Ahmad, N.A.; Abdul, R.I. Gaps between competence and importance of employability skills: Evidence from Malaysia. *High. Educ. Eval. Dev.* **2019**, *13*, 97–112. [CrossRef]
3. Kálmán, A.; Molnár, G.; Szűts, Z. Issues of Lifelong Learning—Behavioral ends of teaching and learning through ICT. In Proceedings of the 9th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Budapest, Hungary, 22–24 August 2018; pp. 395–398. [CrossRef]
4. Stromquist, N.P.; da Costa, R.B. Popular universities: An alternative vision for lifelong learning in Europe. *Int. Rev. Educ.* **2017**, *63*, 725–744. [CrossRef]
5. Van Laar, E.; van Deursen, A.J.A.M.; van Dijk, J.A.G.M.; de Haan, J. The relation between 21st-century skills and digital skills: A systematic literature review. *Comput. Hum. Behav.* **2017**, *72*, 577–588. [CrossRef]
6. Knowles, M.S.; Elwood, F.H., III; Swanson, R.A. *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*; Routledge: London, UK, 2007; pp. 132–139.
7. Laal, M. Lifelong learning: What does it mean? *Procedia Soc. Behav. Sci.* **2011**, *28*, 470–474. [CrossRef]
8. Keinänen, M.; Ursin, J.; Nissinen, K. How to measure students' innovation competences in higher education: Evaluation of an assessment tool in authentic learning environments. *Stud. Educ. Eval.* **2018**, *58*, 30–36. [CrossRef]
9. van Deursen, A.J.A.M.; Verlage, C.; van Laar, E. Social Network Site Skills for Communication Professionals: Conceptualization, Operationalization, and an Empirical Investigation. *IEEE Trans. Prof. Commun.* **2019**, *62*, 43–54. [CrossRef]
10. Melão, N.; Reis, J. Using Social Networks in Personnel Selection: A Survey of Human Resource Professionals. In Proceedings of the 2020 15th Iberian Conference on Information Systems and Technologies (CISTI), Sevilla, Spain, 24–27 June 2020; pp. 1–6. [CrossRef]
11. Ponte, M.C.U.; Zorrilla, A.M.; Ruiz, I.O. Taxonomy-based hybrid recommendation system for lifelong learning to improve professional skills. In Proceedings of the 2020 IEEE International Conference on Engineering, Technology and Education (TALE), Shenzhen, China, 8–11 December 2020.
12. Aggarwal, C. *Recommender Systems. The Textbook*; Springer: New York, NY, USA, 2016.
13. Akmaliah, F.; Krisnadhi, A.A.; Sensuse, D.I.; Rahayu, P.; Wulandari, I.A. Role of Ontology and Machine Learning in Recommender Systems. In Proceedings of the 2018 Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS), Batu, Indonesia, 9–11 October 2018; pp. 371–376. [CrossRef]
14. Shao, J. Research on Fuzzy Ontology E-learning Based on User Profile. In Proceedings of the 2017 International Conference on E-Education, E-Business and E-Technology (ICEBT 2017), Toronto, ON, Canada, 10–12 September 2017; pp. 46–49. [CrossRef]
15. Xu, Z.; Tifrea-Marciuska, O.; Lukasiewicz, T.; Martinez, M.V.; Simari, G.I.; Chen, C. Lightweight Tag-Aware Personalized Recommendation on the Social Web Using Ontological Similarity. *IEEE Access* **2018**, *6*, 35590–35610. [CrossRef]
16. Pereira, C.; Campos, F.; Ströele, V.; David, J.; Braga, R. BROAD-RSI—Educational recommender system using social networks interactions and linked data. *J. Internet Serv. Appl.* **2018**, *9*, 7. [CrossRef]
17. Gil Robles, M. Organizational Transformation during COVID-19. *IEEE Eng. Manag. Rev.* **2020**, *48*, 31–36. [CrossRef]
18. Bakhshinategh, B.; Spanakis, G.; Zaïane, O.; ElAtia, S. A Course Recommender System based on Graduating Attributes. In Proceedings of the 9th International Conference on Computer Supported Education, Porto, Portugal, 21–23 April 2017.
19. Patel, B.; Kakuste, V.; Eirinaki, M. CaPaR: A Career Path Recommendation Framework. In Proceedings of the 2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService), San Francisco, CA, USA, 6–9 April 2017; pp. 23–30. [CrossRef]
20. Bañeres, D.; Conesa, J. eOrient@—A Recommender System toAddress Life-Long Learning and Promote Employability. In Proceedings of the 2016 International Conference on Intelligent Networking and Collaborative Systems (INCoS), Ostrawva, Czech Republic, 7–9 September 2016; pp. 351–356. [CrossRef]

21. Ibrahim, M.; Yang, Y.; Ndzi, D.; Yang, G.; Al-Maliki, M. Ontology-Based Personalized Course Recommendation Framework. *IEEE Access* **2019**, *7*, 5180–5199. [CrossRef]

22. Diaby, M.; Viennet, E. Taxonomy-based job recommender systems on Facebook and LinkedIn profiles. In Proceedings of the 2014 IEEE Eighth International Conference on Research Challenges in Information Science (RCIS), Marrakech, Morocco, 28–30 May 2014; pp. 1–6. [CrossRef]

23. Kuznetsov, S.; Kordík, P.; Řehořek, T.; Dvořák, J.; Kroha, P. Reducing cold start problems in educational recommender systems. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 3143–3149. [CrossRef]

24. Koh, M.; Chew, Y. Intelligent Job Matching with Self-learning Recommendation Engine. *Procedia Manuf.* **2015**, *3*, 1959–1965. [CrossRef]

25. Valentin, C.D.; Emrich, A.; Lahann, J.; Werth, D.; Loos, P. Adaptive Social Media Skills Trainer for Vocational Education and Training: Concept and Implementation of a Recommender System. In Proceedings of the 2015 48th Hawaii International Conference on System Sciences, Kauai, HI, USA, 5–8 January 2015; pp. 1951–1960. [CrossRef]

26. Tarus, J.K.; Niu, Z.; Yousif, A. A hybrid knowledge-based recommender system for e-learning based on ontology and sequential pattern mining. *Future Gener. Comput. Syst.* **2017**, *72*, 37–48. [CrossRef]

27. Harrathi, M.; Touzani, N.; Braham, R. A Hybrid Knowlegde-Based Approach for Recommending Massive Learning Activities. In Proceedings of the 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA), Hammamet, Tunisia, 30 October–3 November 2017. [CrossRef]

28. Cerón-Figueroa, S.; López-Yáñez, I.; Alhalabi, W.; Camacho-Nieto, O.; Villuendas-Rey, Y.; Aldape-Pérez, M.; Yáñez-Márquez, C. Instance-based ontology matching for e-learning material using an associative pattern classifier. *Comput. Hum. Behav.* **2017**, *69*, 218–225. [CrossRef]

29. Balachander, Y.; Moh, T. Ontology Based Similarity for Information Technology Skills. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), Barcelona, Spain, 30 August 2018; pp. 302–305. [CrossRef]

30. Gugnani, V.K.; Kasireddy, R.; Ponnalagu, K. Generating Unified Candidate Skill Graph for Career Path Recommendation. In Proceedings of the 2018 IEEE International Conference on Data Mining Workshops (ICDMW), Singapore, 17–20 November 2018; pp. 328–333. [CrossRef]

31. Obeid, C.; Lahoud, I.; Khoury, H.E.; Champin, P. Ontology-based Recommender System in Higher Education. In Proceedings of the Companion Proceedings of the Web Conference, Lyon, France, 23–27 April 2018.

32. Aissaoui, O.E.; Oughdir, L. A learning style-based Ontology Matching to enhance learning resources recommendation. In Proceedings of the 2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 19–20 March 2020. [CrossRef]

33. Liu, H.; Jiang, G.; Su, L.; Cao, Y.; Diao, F.; Mi, L. Construction of power projects knowledge graph based on graph database Neo4j. In Proceedings of the 2020 International Conference on Computer, Information and Telecommunication Systems (CITS), Hangzhou, China, 5–7 October 2020; pp. 1–4. [CrossRef]

34. Ricci, F.; Rokach, L.; Shapira, B. *Recommender Systems Handbook*, 2nd ed.; Springer: New York, NY, USA, 2015; pp. 265–308.