Recommended system of personalized corporate e-learning based on ontologies

Lyudmila Nikitina¹, Pavel Shikov¹*, Yuriy Shikov¹, Anna Bakanova², and Aleksey Shikov²

¹St. Petersburg State University industrial technologies and design, 191186, St. Petersburg, ul. Bolshaya Morskaya, 18. Russia
²St. Petersburg national research university of information technologies, mechanics and optics, Kronverkskiy pr., 49. Russia

Abstract. Corporate e-learning is becoming general, the use of electronic learning systems is constantly expanding. Corporate training is characterized by the availability of a corporate knowledge base, which is constantly expanding. In order to achieve stable development of company, transparency and access to knowledge and information is necessary. This ensures a thorough understanding of the executed work by all employees of the company. Therefore, a key objective in this context is a generalization, accumulation and transmission of corporate knowledge to any employee of the company under the corporate training. In addition, it is crucial to find the suitable content in the learning process in order to save time and quickly get involved in the workflow. Thus, new systems are needed that can effectively recommend the necessary educational material for users on the basis of their preferences.

1 Introduction

As a step towards providing users with this type of system, an integrated system of recommendations based on ontology is proposed. It is known that the technology of the semantic network effectively manages and represents knowledge. Ontology is used in the semantic network to effectively determine the semantic relationships between learning objects. Ontologies allow one to formally and explicitly define concepts that appear in a particular area, their properties and relationships. Ontology of the trainee and ontology of the subject area of training are used to represent and simulate knowledge about the employee and the field of training respectively.

The success of the company today, on the one hand, depends on the ability to adapt to changes or implement them in the shortest possible time, and on the other hand, to maintain the efficiency of internal processes. This is possible only if the employees have the opportunity to quickly acquire new knowledge and skills. The company, in turn, should be able to retain all the knowledge and experience of employees for further use, and, regardless of whether they are implicit or explicit.

* Corresponding author: pavel.shikov@mail.ru
Ontology of the trainee employee, presents information about the learner and his relationship with the learning concept, for example, «the concept of training is suitable for training», «the concept of training is completed» and «the training course is mastered» [1].

In order to recommend the most suitable content for users, the recommendations methodology combines four basic approaches: matching user profiles, pre-creating learning paths, collaborative filtering, and the method of semantic similarity.

This kind of system successfully solves the problem of lack of information, supporting an ontological approach to profiling users, and, also, significantly improves the accuracy of recommendations. The learner's context here is dynamic and includes previous knowledge, interests, speed and style of learning and goals. Also, the system is constantly being replenished with new implicit knowledge, which employees of the company have [2]. Due to the ontological approach, it is possible to externalize implicit knowledge into the database in an explicit predefined form. The content of the training system and the database of educational materials are being updated and internalization of implicit knowledge by users of the system takes place.

Proper reception, use and management of corporate knowledge is the basis of the efficiency of most enterprises. Knowledge management in the companies of the industry is aimed at preserving and disseminating explicit and implicit knowledge, and knowledge itself is perceived as corporate assets and appropriately valued and protected.

Adding new knowledge to the corporate database, one way or another, occurs through the following stages:
- externalization of knowledge from the source;
- formalization of knowledge in the format adopted by the organization;
- verification of acquired knowledge;
- providing access to new knowledge for employees;
- adding new knowledge to the overall coherent scheme.

The implementation of the stages listed differs in each company, but the common factor remains the coherence of the available knowledge. Connectivity is directly related to the quality of the decisions made, since information torn out of the context can lead to an erroneous conclusion and result in losses for the company. Nevertheless, to make informed decisions, employees of construction companies need a lot of knowledge in their subject area or field of activity. This often leads to a time-consuming process of searching for knowledge necessary to solve production problems or making managerial decisions. The ontological model can accelerate the process of finding knowledge; its semantic links will form the graphs of knowledge. Their use will not only take into account the semantics of search queries, but also create a question-answer system. Even in a large, well-structured knowledge base, it may be challenging to find the necessary information. It is much easier to ask a direct question, a direct answer to which can immediately be used. A number of interesting approaches to the use of ontologies in various learning systems are outlined in many articles [3], [4]. The difference of the approach developed by the authors is the proposal to take into account in the recommendation system personal preferences of employees while forming individual learning paths.

To implement such a system within the framework of corporate training, it will be crucial not only to develop ontological models, but also to make it possible to adjust them. The simplified scheme may be as follows:
- The employee of the company formalizes new knowledge in accordance with the format adopted by the company.
- The system automatically creates a semantic model of new knowledge.
- The system builds a model of new knowledge into the general ontology of the company's knowledge.
In this system, two ontologies can be considered basic: the ontology of the student and the ontology of the subject domain. The first one is used for representation, modeling and retrieval of knowledge about the company's employees, the second one, for searching the company's business information.

Ontologies take the most important place in the systems of electronic education, where individual trajectories and personalization of education have become the most important aspects. Ontological models are widely used in the field of corporate training, both in Russia and abroad, including development of individual educational paths in order to organize personalized training and professional development.

2 Methods and models

2.1 Ontological model for the formation of teaching materials in corporate e-learning systems

This problem can be solved by a specific algorithm. Developing a list of tasks for different positions in the company and describing the solutions will significantly shorten both the time for performing these tasks and the quality of their results. However, not all solutions of work tasks can be unambiguously formalized. Some even require complex decisions in conditions of uncertainty. In such cases, it is possible to develop and provide employees with an action program that will be a key element of the decision-making workflow. The employee action program is prepared for a specific situation from the workflow, which requires decision making, and includes in the hierarchically structured form of all the necessary knowledge, regulations and corporate rules. The internal structural units of the action program can be called artifacts, examples of which can be:
1. links to websites;
2. notes with the results of meetings and negotiations;
3. emails;
4. text documents;
5. spreadsheets;
6. links to articles in the corporate knowledge base.

As an implementation of action programs could serve a separate module of the information system, in which the user is given the opportunity to add heterogeneous artifacts, group them and save them for later use. In this case, any employee can organize artifacts by placing them, for example, in separate workspaces, and then group or link them together with semantic links. The combined array of created action programs creates another ontology that will allow you to search for solutions to specific problems, as well as provide recommendations on the use of adjacent ones, if required by the prevailing situation. An example is a B2B transaction for the sale of a complex information system. At the beginning of the transaction, the buyer has his own expectations from the future system, which are accounted for by the supplier. The system delivered to the client meets these requirements, but its technical implementation does not allow to integrate without modifications with the existing customer system. Depending on the specifics, the development of the situation may lead either to the cancellation of the transaction, or to the completion of the supplier's system, or to the completion of the customer's system. Each of these options means additional costs for the parties, therefore it is necessary to take a large number of correct decisions so that both sides achieve the expected goals with minimal additional costs. This situation mainly requires a huge amount of knowledge for negotiating, the algorithm of which is almost impossible to plan. Employees of the company are given the opportunity to create a library
of solutions for all cases - working cases, as well as the possibility of their search and actualization.

The case creation cycle begins with a description of a new problem in the system that converts this description into a representation of a particular case using the situational model definition mechanism and has four main steps:

1. obtaining the most similar scenarios for the use of cases with the most similar description of the problem;
2. choice of the closest case to the situation being solved;
3. reuse of the extracted case;
4. if the result is unsatisfactory, then the case is reviewed, and either its correction or the creation of a new one should be done;
5. testing a new case or a revised case;
6. preservation of changes or a new case in the solution library (corporate knowledge base).

An illustration of this kind of cycle is shown in Figure 1. However, the construction of such a model of work with cases, in the first place, requires the presence of an ontology reflecting the specificity of a particular subject area.

![Fig. 1. Cycle of work with corporate cases.](image)

The accumulated library of cases is an additional means of forming complexes of scientific materials for employees. In fact, each case is a completed training block containing all the necessary information that an employee can study and then complete the task. Each position in the company can be described as a set of knowledge, skills and abilities that a candidate must possess in order to occupy it. In addition, the ontology made up of them will allow us to find similar ones, and then group them into special modules to categorize the internal tasks of employees. Their study will provide a much deeper understanding of the company's internal processes, as well as offer or adopt best practices by solving relevant cases.
2.2 Case Model

All existing cases in the company must undergo processing and get additional attributes and meta-information in order to be added to the general ontology. The main processing procedure is the conceptual transformation of data. Its main goal is to find words or phrases in artifacts of a certain case that describe the concepts existing in the corporate domain ontology. For this, the components, called conceptual converters, use dictionaries linking the unique name of the concept, its identifier with different spelling variants. Due to the processes of tokenization, the removal of stop words (for example, "is", "on", "etc."), decreasing the register and morphological search for input texts and dictionary entries, the conceptual data transformation is able to find more different spellings of the word, than in the usual dictionary. At the concept presentation stage, each artifact in the case is processed through a pipeline containing conceptual maps associated with the dictionaries (Figure 2).

![Fig. 2. Creation of a case model.](image)

At the end of this step, a model with several sets of concepts is created, each of which is associated with a specific specific type. In this process, the text is first extracted from artifacts, then processed using a conceptual converter associated with the vocabulary of concepts, and then various sets of concepts are combined with the case model.

Due to the introduction of ontological preliminary processing of data, it is possible to achieve three goals:

- initially limit cases to the most important concepts for a specific area and measure the similarity only for the essential parts of the case;
- to achieve a higher similarity of the conceptual level from the point of view of the correctness of the term (that is, to exclude the polysemy of words);
- The extracted ontological concepts can be used to analyze semantic similarity in terms of evaluating the semantic similarity.

Ontologies are very often used for other purposes. For example, to compare ontologies, their characteristics and assess the quality of corporate e-learning resources. There are a lot of tools for working with ontologies, which allow to analyze their characteristics and compare several ontologies among themselves.

2.3 Personalization of corporate training

In the field of corporate training, the aspect of personalized training is extremely important for the development of professional competences of personnel. Modern views on these
processes have been studied in a number of publications [5], [6], [7]. Depending on preferences, inclinations and predispositions, employees want to develop either as part of their specialization, that is, to expand their terms of reference while remaining in their posts, either to engage in highly specialized work or to take a managerial position.

In this case, there is a horizontal or vertical development of the career and, depending on the chosen path, different training is necessary. Accordingly, the employee needs training on a certain trajectory, which will allow him to get all the knowledge, skills and skills necessary for a new position or activity.

Given all the ontologies mentioned above, it is logical to develop one more for posts in the company. It will allow you to identify similar posts, identify hidden links between, at first glance, independent, and also explicitly link posts with the necessary knowledge, abilities and skills.

To build an individual trajectory of learning, you must use the ontology of the positions of the company, as well as the ontology of the subject area. The first ontology searches for and creates a list of differences between the current position of the employee and the one for which he aspires, and the second ontology will form the thematic content of the training course, with the passage of which the existing difference in the competences of the employee will be eliminated.

An important aspect is the possibility of using semantic dependencies in ontologies, which can significantly increase the complexity and depth of training.

Moreover, for each specific employee, an individual list of recommended posts can be drawn up, which he can later occupy. The basis of the recommendation may be semantically related cases of recommended posts, the same knowledge, skills and abilities.

### 2.4 Competence profile of company employees

The user profile is a key component of the corporate training system. It stores the history of the employee's training, the list of competencies that he owns, as well as a description of his current position and all previous ones (if they ever existed).

The profile of the competencies of each employee is naturally related to the ontology of the subject area through the knowledge that it possesses and the cases that it solves, and it is also related to the ontology of posts, through the posts that the employee held in the past or currently occupies. There can be a situation when the employee will have an unconventional set of knowledge, skills and abilities, since he held various positions or performed various tasks. This kind of employee is especially valuable for the company, as they have extensive experience and know the company's activities from different sides.

Training in the company can also be based on a scheme in which one employee trains another, structuring and sharing their knowledge, as well as practical experience. For this, it is necessary to ensure the search for the necessary knowledge and experience in the context of employees. Available ontologies of positions, knowledge, skills and abilities will undoubtedly allow this. Furthermore, the profile should display the history of the generated recommendations for the employee, both in terms of studying the competencies of the current position, and potential ones.

In addition to the already mentioned ontologies, it is useful to use one more, dedicated to the training activities of the company. It must explicitly link the training courses, their authors, the positions for which they are intended, cases that will complement the core training, as well as competences. An example of a part of such an ontology is shown in Figure 3 and represents information about the learner and his relationship with the elements of learning.

The attitude "requires" the element of "learning with respect to oneself" and describes the existence of an "element of learning" that must be before the next element is studied. That is,
for the training, an employee must already have some knowledge to select a specific element of training in the future. Similarly, the relation "has URL" is described.

Fig. 3. Example of learning ontology.

An important aspect of providing a context-sensitive learning function is a recommendation system that provides the user with advice on the content, taking into account the context of the learner. These recommendations help to find the right content and improve the quality and speed of training.

This sort of system recommends studying content based on four main approaches:
1. The approach based on the user profile; includes recommendations on the content of information based on data collected about the user and his preferences.
2. The recommendation on training requires preliminary recommendations, which must be studied before taking the specific content of the course. Each concept is recursively requested to create a structured learning path in accordance with the preliminary concepts and then delivered to the student.
3. A recommendation based on semantic similarity offers content that is similar to what has been learned at the moment.
   Each of these four approaches uses knowledge and ontology data.
   The user profile is a key component of the recommendation system. It tracks the current user and his actions for e-learning. The system implements the method of social learning, which includes subsequent topics of interest [8]. The user profiling system works in two directions. The explicit actions of the profile system are form-based components that the user
represents as the designation of the completed course, following the content topics and following other users. Implicit actions in the profile system include the automatic saving of user actions, such as the generation of a list of recommendations dynamically based on the history of user viewing, the calculation of the number of sub-concepts within the framework of the basic concept, and the dynamic generation of a tree-like representation of all concepts.

As in conventional curricula, each course can have another mandatory course as a condition for passing. Therefore, for each concept in an ontology, you need to define a property that contains concepts that you need to study before the course begins.

It is important to note that the preliminary recommendations may have their own necessary conditions for study. Therefore, for correct interaction of the user with the learning process, a path is needed.

Due to recursive requests for the identification of preconditions for a particular element of learning, it is possible to search for them until the most basic concept of the learning element is reached, which has no preliminary recommendations. After this, you can build a whole path from the base element to the required one, which will be called the training trajectory. The algorithm for constructing this training path must ensure that it does not contain cycles and creates a topological structure of courses based on a deep search for concepts of existing subject ontology.

A Recommendation based on semantic similarity uses similarity methods and similarity concepts for linking different learning objects, that is, fragments of educational material. The algorithm is based on a method that calculates the semantic similarity between learning objects. It works by constructing a semantic map of the material of e-learning on the basis of calculated estimates of semantic similarity. These estimates are calculated taking into account the frequency of occurrence of the connection of words in the case. The first step of the algorithm is to remove noise from the text (common words, which are insignificant).

The remaining words are target, and are stored in the table by a value that depends on the keyword. Target words in different tables are compared, and this procedure is performed through the entire text.

The system of recommendations is a step-by-step definition of the training conditions necessary for the employee. The history of training shows the general specialization of the employee, determines the knowledge that he has as a result of previous training. However, many skills are acquired in the process of work, therefore, the set of competences that the employee already possesses is taken into account. Interests determine what competencies and skills he lacks for further work on the direction of his interest.

The learning style and user profile specify possible learning conditions. All these initial and very important recommendations are combined and, based on the key words (semantic map) and similarities of parameters, definitions, which the employee already knows, possible training courses are determined. At this stage, specialization of training (narrow direction or universal) is also determined. Further in this specialization the recommendations of employees, similarity of various users and their results in the chosen program are taken into account. Only after that new content is added, that is, specific topics for study.

Preliminary recommendations include the formation of a path, and the agreement with the terms of the course. For example, the terms of the training, the type of training depending on the distance (remote or full-time), the type of equipment required, etc.

As more and more content becomes available on the Internet, the need to provide relevant and personalized advice to learners is becoming increasingly crucial. The article suggests an integrated semantic advisory approach using ontology to recommend training material for relevant and personalized educational content.

3 Conclusions and recommendations
The presented advisory system based on ontologies is able to offer relevant information for an experienced specialist who processes sophisticated tasks based on similar completed cases stored in the database of the system. The system of recommendations extracts a model specific to the subject area (based on the provided ontology of the domain), which is used to assess the similarity between the cases. The information offered by the system is determined by the priority of its relevance for the current course within the specialist's work environment. Thus, a new model of ontological search for similar cases is presented, based on the text of the prioritization of artifacts for the selected topic and a specific business problem. Moreover, the system of recommendations also takes into account the user profile and the joint filtering, that is, information or any other recommendations of previous or current employees of the company. Within the framework of the conducted researches the general architecture of the recommendations system was developed and presented.

In the future, it is planned to add additional criteria for the system of recommendations, such as training time and additional filtering recommendations based on the style of instruction, knowledge indicators, a set of skills and competencies. The proposed training trajectories can also be added on the basis of previous experience of past users of the system.

References