

**WELLSO 2017**  
**IV International Scientific Symposium Lifelong Wellbeing in  
the World**

**THE NEW ALGORITHMS OF MACHINE LEARNING FOR  
EDUCATION PEOPLE WITH SPECIAL NEEDS**

Alena Khaperskaya (a)\*, Olga Berestneva (b)

\*Corresponding author

(a) Tomsk Polytechnic University, 634050 Lenin Avenue, 30, Tomsk, Russia, khape@mail.ru, +79521799945

(b) Tomsk Polytechnic University, 634050 Lenin Avenue, 30, Tomsk, Russia, ogb6@yandex.ru, +79131061994

*Abstract*

The paper addresses the problem of adapting people with special needs to their environment. We support the idea that organizing special algorithms will be a solution to this problem. The analysis of people behavior in the actual learning process and their e-learning experience shows their ability to adjust their actions and develop adaptation skills relevant to any environment. Furthermore, we analyze the ways to involve people with special needs into the virtual setting activities, thus enabling them to feel that they are productive employees and members of the society. We present the detailed algorithm of the intellectual research, where each step affects the overall decision-making process. Participants, including those with special needs, can also correct their decisions, which helps them develop their abilities to adapt to their future working environment in a company. The main advantage of arranging such process in the electronic environment is that people with special needs acquire the adaptation, communication and decision-making skills as part of machine learning. The analysis of the subject area was carried out, and the main problems of creating automated systems for searching competency development tasks were considered. Also the methods that are used for reference systems (collaborative filtering), information semantic search, and separation of texts on topics without training are presented.

© 2018 Published by Future Academy [www.FutureAcademy.org.UK](http://www.FutureAcademy.org.UK)

**Keywords:** Semantic analysis, disable people, competences, lsa-algorithm, data mining, machine learning.



This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 Unported License, permitting all non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## 1. Introduction

The present-day system of professional education does not only aim to improve the quality of professional training and competitive strength of university graduates but also to form new ways of thinking and intellect adapted to fast-changing technological, social, economic and information environment. When someone enters a working environment for the first time, they often have to face conflicts in the organization, solve complex problems, work in a team, be a leader and make important decisions that will affect the future of the organization as well as consider other social aspects. Not everyone, however, knows how to act in these situations. This problem can be solved by arranging business games in a virtual machine. This type of activity is also gaining popularity as an effective supplementary form of social support for people with special needs. Virtual business games make it possible for all people including those who are mobility challenged to actualize their economic, political and creative subjectness.

A potentially useful intelligent tool for researchers is the opportunity to establish the connection between the facts which aren't connected with each other, and, as a result, the new ideas, approaches and hypotheses appear. This process is known as the analysis of the text (or intellectual data analysis if it is aimed at not bibliographic data set).

Data mining (text mining) includes the analysis of the big collection of the digital elements, which are often not connected with each other on a systematic basis and there is the process of the unknowns' detection before the facts, which could serve further as templates. If being guided by traditional manual search it would be extremely difficult to find complex couplings between the objects.

Now the ability of text recognition and data mining is the accelerator of the research process and the effective operation which was done in the past. Nevertheless, a row of features or conditions arises before starting such type of activity (Dekelver et al., 2015). These include:

- access to a huge set of information on scientific research in the sequential and compatible form;
- free access without prohibitive monitoring of authentication;
- the digitized text, data and other media sources;
- data which aren't protected by copyright, etc.

The diversity, value and speed of data increase every year (Saltykov et al., 2015). To track and process the changing data these hardware opportunities also should undergo essential changes.

As data mining is not a single operation, but a set of steps, theories and algorithms, the hardware can be partitioned into a row of components. Changes in components aren't always synchronous with the increase in demand in data mining, machine training and big analytical problems solving. A hardware system of data mining must be balanced in relation to components to give users the best solution of their analytical problem (Gallant, 1988).

Data mining of any scale can't be made without specialized software or virtual machine.

## 2. Problem Statement

Social, psychological and professional adaptation is one of key importance for successful work in an organization. This involves a wide range of topical problems under investigation (Berestneva, et al, 2015a). These types of adaptation have a complex hierarchy and their own special dynamics of intercorrelation at different stages of learning. In the end, however, they integrate into a single adaptation process, in which the adaptation medium and the adapting element (disable people) interact to engender the adaptive situation encouraging the adaptive need in students.

The theoretical analysis and practical investigation into the adaptation factors of people with special needs helped us identify the organizational and pedagogical conditions of efficiency of each adaptation type (Berestneva, et al, 2015b):

1. Forming cognitive motivation and professional activity of disable people with due consideration of the virtual environment of the teaching process;
2. Forming a disable people group with a favorable social and psychological climate; psychological comfortable, moral satisfaction of students with themselves and with their teamwork;
3. Forming a steady focus on acquiring professional qualification.

To implement the first condition, we need to center around the systematic forming of students' mental efforts and the approach to differentiating the levels of learning and cognitive activities of disable people. With this in mind, the first stage of the adaptive process incorporates the formation of cognitive activity at the reproductive level; the second stage is devoted to the heuristic level and the third one focuses on the creative level.

Since it is important to consider the nature of the computer information environment of the learning process, when forming the cognitive activity of disable people, the first pedagogical condition is provided by purpose-designed virtual business games. There are detailed methodology guidelines uploaded to the university's corporate portal for both teachers and students. Thus, students have free access to the whole set of cutting-edge learning tools including those for distance learning and virtual training.

The implementation of the second pedagogical condition of social and psychological adaptation for disable people is associated with the following (Berestneva, et al, 2015c):

1. Arranging teamwork for people with special needs a favorable social and psychological climate through involving them in active creative cooperation;
2. Providing psychological comfort as well as moral satisfaction of students with themselves and being a part of the team (Bermus, 2005).

The leading role in this type of adaptation falls on supervisors and dean's office representatives responsible for discipline as well as subject teachers.

The third pedagogical condition — providing a stable focus on the acquisition of professional qualification (professional adaptation) — is especially important for further development of professional competences. A stable focus on professional qualification comes from positive attitude to the profession as well as features of character that are professionally in demand. The implementation of the above pedagogical conditions will result in the development of skills listed in Fig 1. It also outlines the outcomes of developing this or that skill.

In this paper, we address the problem of disable people adaptation to their future working environment in a company and offer ways to solve it in the form of virtual games and various business cases in the online setting. This field is also supplemented by the analysis and integration of people with special needs to the virtual environment. Thereby, giving them the opportunity to feel the real organization as a full unit of society. We have also identified the pedagogical conditions of social and psychological adaptation of disable people and people with special needs. The research has shown that the provision of these conditions will result in the development of skills specified at Fig. 01.

Thus, despite the active research into the impact of the global network on various sides of the society, the investigation into the influence of business games in an electronic setting on social adaptation of students and physically challenged people is still largely overlooked by most scientists.

As the key evidence, we are going to give an example of a business case and intellectual search algorithm. We have also identified the pedagogical conditions of social and psychological adaptation of people with special needs. The research has shown that the provision of these conditions will result in the development of skills specified at Fig 01.

Thus, despite the active research into the impact of the global network on various sides of the society, the investigation into the influence of business games in an electronic setting on social adaptation of physically challenged people is still largely overlooked by most scientists (Khaperskaya, et al., 2016).

<b>Individual skills</b>	
Being able to single out specific objective crucial for achieving the main goals.	Forming independent judgment
Developing personal interest	Acquiring self-organization skills
Deepening the knowledge in a certain sphere	
Seeing the prospects of developing practical skills	Creating something individual
Performing activities that are of practical use	Creating something unique
Developing the ability to adapt to critical situations	Collecting and analyzing new information
Explore problem solving strategies (algorithms) and integrate them into practice	Analyzing and evaluating other people's work
Learning to motivate yourself at work	Developing personal proactivity
<b>Interdisciplinary skills</b>	
Applying one's knowledge beyond a specific field	Integrating the knowledge acquired from different sources
Learning to perceive different situations, approach them from different standpoints and be able to adapt to them	Being able to perceive criticism and being ready for contradictory statements and arguments.
<b>Teamwork skills</b>	
Learning to work in a team	Being able to be a team leader and having organization skills
Developing diplomacy and tact	Supervising and guiding other people's activity
Taking part in the decision-making process	Acquiring cooperation skills
<b>Work with self-awareness</b>	
Doing one's own SWOT analysis (strengths, weaknesses, opportunities, and threats)	Acquiring satisfaction with the job done
Objectively estimating one's abilities in terms of solving the set objective	Striving for the feeling of autonomy and freedom in one's professional field
<b>Communication skills</b>	
Improving the perception of information the electronic format and formulation of questions while digesting this information	Writing a comprehensive and clear progress report
Offering one's work to other people's criticism (both written and oral)	Improving leadership and persuasion skills following a logical line of reasoning

**Figure 01.** Skills acquired by disable people when doing a business case on the virtual machine

### **3. Research Questions**

The relevance of this work is the development of algorithms for intellectual search and software to search for assignments of competence development for people with special needs. The research questions were solved to achieve this goal:

- The general approaches to information retrieval are revealed;
- The modern search systems are analyzed and the main shortcomings are revealed;
- The expediency of automating the process of searching for assignments for the development of competencies is justified;
- The algorithm of intellectual search is developed;
- The developed unique adaptation program that allows people with special needs to search for tasks for the development of competencies.

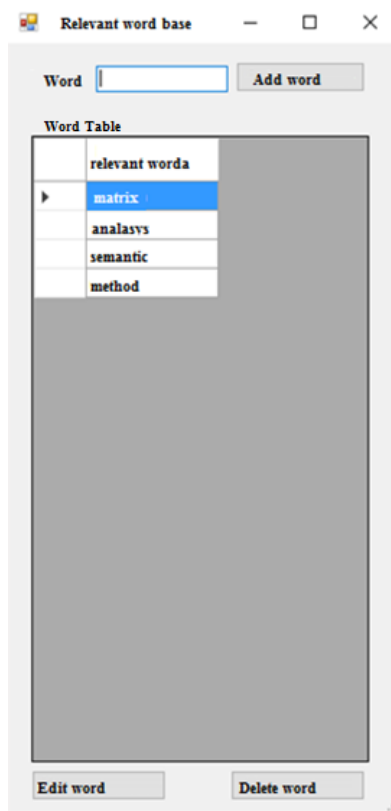
### **4. Purpose of the Study**

The purpose of the study is determined by:

- the creation of a competent approach in the implementation of the search for the development of competences and data processing, due to the large number of information and mechanisms for processing the results;
- the reduction of the trainings cost and development programs for the development of competencies;
- the creation of adaptive software for the development of competencies for people with disabilities in health;
- the reducing the burden on the teaching staff.

### **5. Research Methods**

On the basis of the intelligent data processing methods, we can create a semantic core that contains all the keywords associated with a particular competence. Figure 2 shows the interface result, which contains the relevant words database, where the user can enter his query. Then, the task will be identified or recognized on the basis of the entered competence. After studying and passing the mentioned tasks the user will be able to develop this competence.



**Figure 02.** Relevant words base for users

This program, which shows a Relevant word base for users, was developed in Matlab.

As an example, we can consider this kind of competence to be a communication activity, as well as rapid decision-making. Table 2 shows the examples of two competencies, as well as a few words related to the development of these competencies, which can be found in a business game, business case, problem situation, etc. All the keywords related to this concept are presented in Table 01.

**Table 01.** Set of words for creating a semantic core of a certain competence

<b>Communication activity</b>	<b>Rapid decision-making</b>
Conversation	Reflections
Decision	Analysis of the situation
Teamwork	Reasonable arguments
Teambuilding	Previous experience
Gestures	Improve the process
Mimicry	Efficiency
The words	Definition
Negotiating skills	Clarity
Speeches	Speed

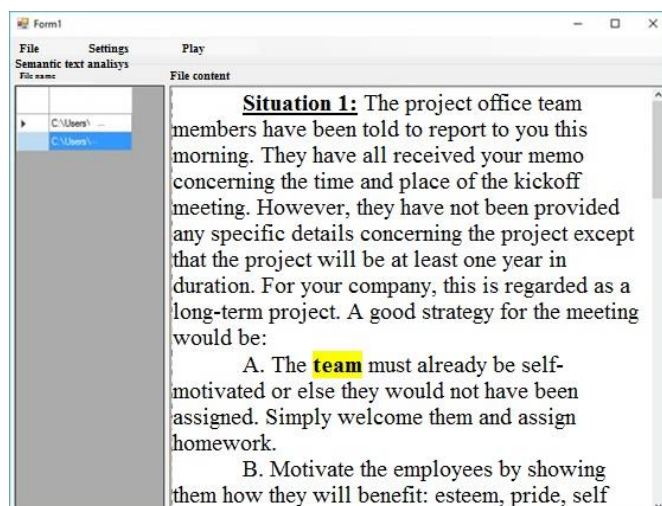
According to the competence, such as communication activity, entered in the search, several business cases containing binders with this word concept were issued. On a rather small fragment of the business case, it can be seen that the program found a binding word, for example - "negotiating skills" and issued the task, after studying which the user can develop communication activity.

That is, general adapted software has been created to search for competency development tasks. The program produces a semantic analysis of the documents that are on the server, in order to find this or that information. Also the program parses the text inside the document and compares the request to it. In this way, we can create a semantic core. When we add a specific document or a task to the core, it is indexed; the program builds a semantic image by some criteria based on the information in the header and some sections, identifies the keywords and uses them to form a tree cloud which highlights the semantic core of the document. The query can be formed in a natural form and then, when it is converted into a semantic tree, it is compared with the database. If the inclusion occurs, certain documents that meet the request will be issued.

The first step in our proposed algorithm is to exclude the stop symbols, that is, words that occur in each text, but do not carry any meaning, for example, all conjunctions, particles, prepositions and many other words.

The next step in the algorithm for creating automated task search for the development of competencies is the need for the operation of a stamping operation. It is not mandatory, it has been experimentally confirmed that good results are obtained without it. In addition, if the set of texts is large enough, then this step can be skipped. The Porter algorithm was used for stamping.

The final step of the algorithm for automated search of tasks for the development of competencies will be the exclusion of words that occur in a single case. This step is also optional; it does not have any significant effect on the final result, but greatly simplifies mathematical calculations. As a result, after all the steps are done, the indexed (relevant) words remain. In the figure they are shown in bold type (Fig.03).



**Figure 03.** An example of searching for a text fragment using the LSA algorithm

The figure presents the result of the search that was created using the proposed algorithm, as well as the LSA algorithm applied to detect hidden patterns or non-obvious dependencies.

Among a huge number of algorithms that are used to search and analyze information, a special place is occupied by those of them whose purpose is to detect hidden patterns or non-obvious dependencies (Berestneva et al, 2015a).

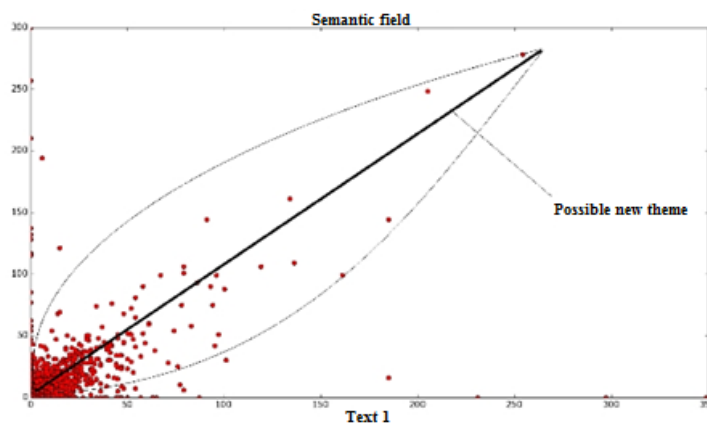
Using them, we can say, for example, that two texts are similar, even if this similarity is expressed indirectly. Or, for example, the words "skis" and "car" belong to different categories, but being used together they can be interpreted in such categories as "sports" and "rest".

One of such kinds of methods used for recommendation systems (collaborative filtering), information semantic search, the separation of texts on subjects without training and many others will be described later. This method is called latent-semantic analysis (LSA-Latent semantic analysis).

But in real life there are a lot of documents and words and there arise the following problems:

- dimensionality (computation of proximity between vectors becomes a slow procedure);
- noisiness (for example, small foreign text insertions should not affect the subject m);
- low level (most cells in the table will be zero ones).

In this case, the following idea is quite logical: instead of using the word-document table, the term-subject and the subject-document should be applied. The solution of this problem is offered by the LSA. But, unfortunately, the interpretation of the results can be difficult.



**Figure 04.** Example of two cards of business cases for competence development

Figure 04 shows an example of two cards of business cases for competence development. It is seen that they have individual and common characteristics, and then a new topic appears.

LSA limitations are as follows:

- It is impossible to get more subjects than documents or words.
- The semantic meaning of a document is determined by a set of words that, as a rule, go together.
- Documents are viewed as simple sets of words. The order of words in documents is ignored. It is only important how many times a word occurs in the document.
- Every word has a single meaning.
  - A disadvantage of the LSA is the assumption that the word map in the documents does not have the form of a normal distribution. Other modifications of the method cope with this problem (Saltykov et al., 2015).

The LSA includes the following steps:

- The process of stop words removing, stamping or lemmatization of words in documents;
- The exclusion of words that occur in a single copy;



- Matrix word-document building (binary existence or absence of a word, the number of occurrences or tf-idf);
- The expansion of the matrix by SVD ( $A = U * V * WT$ ). According to the competence, entered in the search, such as communication activity, several business cases containing binders with this word concept were issued.

## 6. Findings

The findings of the study is determined by:

- The search algorithm, based on the method of data recognition directly in the document itself, containing tasks for the development of specific competencies, different from existing search systems what allows people with special needs accurately to identify the required task, taking into account the rejection of files marked as spam.
- An integrated system of semantic analysis was developed and implemented. The system is used to solve the problems of extracting definitions and relevant terms from texts containing tasks for the development of competences.
- It has been experimentally shown that using the integrated system of semantic analysis significantly improves the accuracy of the search in comparison with the separate application of methods of semantic analysis.
- It has been experimentally shown that the use of SVD-decomposition in the implementation of the LSA-algorithm of matrices increases the efficiency of constructing rules for extracting definitions and relevant terms from texts containing tasks for the development of competencies.

## 7. Conclusion

The task of competencies development has been solved in this work with the help of machine learning. It is relevant for the system of general and professional education. The relevance of the study is determined by the need for further training and skills development, as well as the introduction of information technologies in the education process. The solution of this problem allows improving the methods and ways of developing competences with the help of an automated search system.

The analysis of the subject area was carried out, and the main problems of creating automated systems for searching competency development tasks were considered.

The methods that are used for reference systems (collaborative filtering), information semantic search, and separation of texts on topics without training are presented. Application of the method of latent-semantic analysis (LSA) is described. The developed algorithm can be used in other areas that require text analysis, for example, in search of plagiarism testing systems. The application of the developed algorithm in the system of antiplagiarism will be more effective than the widespread method of "shingles", a significant disadvantage of which is the lack of the possibility of processing synonyms. Using text filtering, stamping and character transformation in the proposed algorithm allows finding

borrowed texts even with their slight modification. Developed software can be recommended for any category of people.

## References

- Berestneva, O. G., Marukhina, O. V., Benson, G. F., Zharkova, O. S. (2015a). Students' competence assessment methods. *Procedia - Social and Behavioral Sciences*, 166, 296-302. DOI: 10.1016/j.sbspro.2014.12.527
- Berestneva, O. G., Marukhina, O. V., Fisochenko, O. N., Romanchukov, S. V. (2015b) Information technology assessment of competence of technical university students. In *International Siberian Conference on Control and Communications (SIBCON)*, IEEE Russia Siberia Section, Novosibirsk.
- Berestneva, O.G., Marukhina, O.V., Kozlova, N.V., Lombardo, C. (2015c). Modelling coping strategies of technical university students. *Communications in Computer and Information Science*, 535, 81-90. DOI: 10.1007/978-3-319-23766-4\_6
- Bermus, A .G. (2005, September) Problems and prospects for implementing a competence approach in education. Retrieved June & July, 2017 [www.eidos.ru/journal/2005/0910-12.htm](http://www.eidos.ru/journal/2005/0910-12.htm),
- Dekelver, J, Kultsova, M., Shabalina, O., Borblik, J., Pidoprigora, A., Romanenko, R. (2015). Design of Mobile Applications for People with Intellectual Disabilities. In *Creativity in Intelligent Technologies and Data Science*. Retrieved June & July, 2017. DOI: 10.1007/978-3-319-23766-4
- Gallant, S. I. (1988). Connectionist expert system. *Communications of the ACM*, 31(2),152-169. DOI: 10.1145/42372.42377
- Khaperskaya, A., Zabrodina, I., Lutoshkina, O., Slesarenko, I., & Bogdanova, A. (2016). Creating a Virtual Enterprise as Part of a Business Game for Adaptation of Students and People with Special Needs. In *8th International Conference on Education and New Learning Technologies*. Retrieved June & July, 2017. DOI:[dx.doi.org/10.21125/edulearn.2016.2389](http://dx.doi.org/10.21125/edulearn.2016.2389)
- Saltykov, S., Rusyaeva, E., Kravets, A. (2015, January). Typology of Scientific Constructions as an Instrument of Conceptual Creativity. In *International Conference on Information, Intelligence, Systems and Applications*. Retrieved June & July, 2017. DOI: 10.1007/978-3-319-23766-4\_4