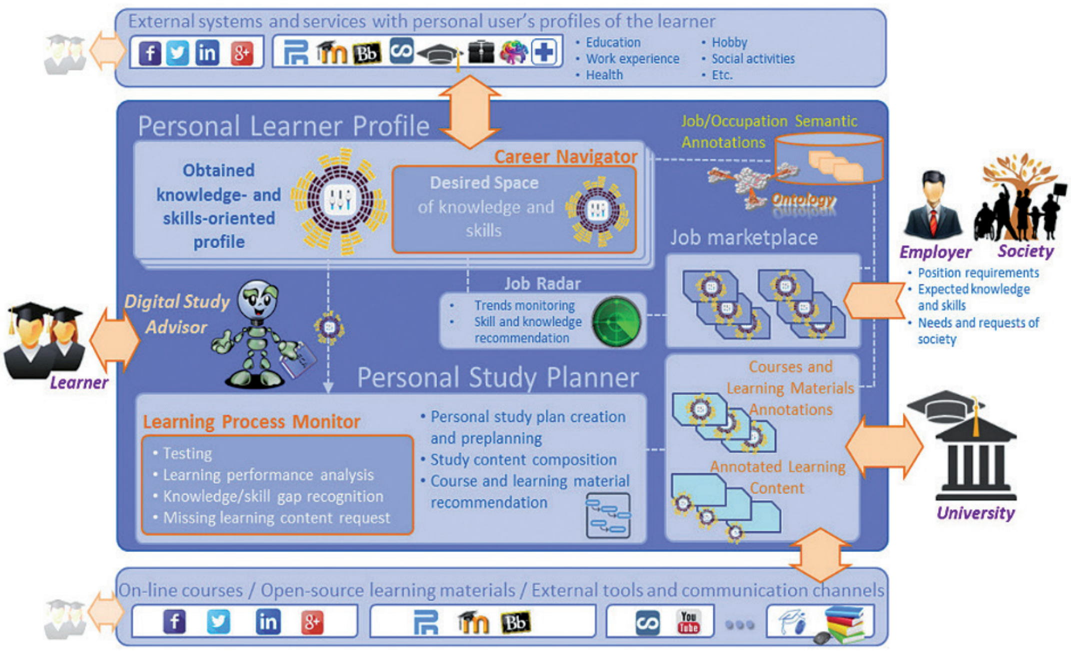
JYVÄSKYLÄ STUDIES IN COMPUTING
272

Mariia Gavriushenko

On Personalized Adaptation of Learning Environments



JYVÄSKYLÄN YLIOPISTO

JYVÄSKYLÄ STUDIES IN COMPUTING 272

Mariia Gavriushenko

On Personalized Adaptation of Learning Environments

Esitetään Jyväskylän yliopiston informaatioteknologian tiedekunnan suostumuksella
julkisesti tarkastettavaksi yliopiston Agora-rakennuksen Alfa-salissa
joulukuun 19. päivänä 2017 kello 12.

Academic dissertation to be publicly discussed, by permission of
the Faculty of Information Technology of the University of Jyväskylä,
in building Agora, Alfa hall, on December 19, 2017 at 12 o'clock noon.

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

On Personalized Adaptation of Learning Environments

JYVÄSKYLÄ STUDIES IN COMPUTING 272

Mariia Gavriushenko

On Personalized Adaptation of Learning Environments

UNIVERSITY OF JYVÄSKYLÄ

JYVÄSKYLÄ 2017

Editors
Timo Männikkö
Faculty of Information Technology, University of Jyväskylä
Pekka Olsbo, Ville Korkiakangas
Publishing Unit, University Library of Jyväskylä

Cover picture by Mariia Gavriushenko and Oleksiy Khriyenko.

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7287-5

URN:ISBN:978-951-39-7287-5
ISBN 978-951-39-7287-5 (PDF)

ISBN 978-951-39-7286-8 (nid.)
ISSN 1456-5390

Jyväskylä University Printing House, Jyväskylä 2017

ABSTRACT

Gavriushenko, Mariia
On Personalized Adaptation of Learning Environments
Jyväskylä: University of Jyväskylä, 2017, 60 p.(+included articles)
(Jyväskylä Studies in Computing
ISSN 1456-5390; 272)
ISBN 978-951-39-7286-8 (nid.)
ISBN 978-951-39-7287-5 (PDF)
Finnish summary
Diss.

This work is devoted to the development of personalized training systems. A major problem in learning environments is applying the same approach to all students: i.e., teaching materials, time for their mastering, and a training program that is designed in the same way for everyone. Although, each student is individual, has his own skills, ability to assimilate the material, his preferences and other. Recently, recommendation systems, of which the system of personalized learning is a part, have become widespread in the learning environments. On the one hand, this shift is due to mathematical approaches, such as machine learning and data mining, that are used in such systems while, on the other hand, the requirements of technological standards "validated" by the World Wide Web Consortium (W3C). According to this symbiosis of mathematical methods and advanced technologies, it is possible to implement a system that has several advantages: identifying current skill levels, building individual learning trajectories, tracking progress, and recommending relevant learning material.

The analysis of feedback, academic advising, and recommendation systems underlies the proposed idea. The conducted research demonstrates how to make learning environments more adaptive to the users according to their knowledge base, behavior, preferences, and abilities. In this research, a model of a learning ecosystem based on the knowledge and skills annotations is presented. This model is a general model of the lifelong learning process. Second, this thesis focuses on the creation of tools for personalized assessment, recommendation, and advising. Third, it is concentrated on developing an adaptive learning game for children, which takes into account the differing perception of words by students during training.

Keywords: personalized learning, adaptive learning, learning environment, Semantic Web, recommendation system, academic advising

Author	Mariia Gavriushenko Faculty of Information Technology University of Jyväskylä Jyväskylä, Finland
Supervisors	Professor Dr. Pekka Neittaanmäki Faculty of Information Technology University of Jyväskylä Jyväskylä, Finland
	Professor Dr. Tommi Kärkkäinen Faculty of Information Technology University of Jyväskylä Jyväskylä, Finland
	Dr. Oleksiy Khriyenko Faculty of Information Technology University of Jyväskylä Jyväskylä, Finland
	Professor Dr. Marja Kankaanranta Faculty of Information Technology University of Jyväskylä Jyväskylä, Finland
Reviewers	Assistant Professor Yuji Tou Tokyo Institute of Technology Tokyo, Japan
	Associate Professor Francesco Guerra University of Modena and Reggio Emilia Modena, Italy
Opponent	Academy Research Fellow Kristian Kiili Tampere University of Technology Tampere, Finland

ACKNOWLEDGEMENTS

While I am writing this text, my thesis is ready and I am happy to be at the final stage of my Ph.D. studies.

I would like to express my deepest appreciation to my Ph.D. Supervisor Prof. Pekka Neittaanmäki for giving me a possibility to work for the university, for believing in me, for supporting and guiding me throughout this research work.

In addition, my deepest appreciation goes to my co-supervisors Prof. Tommi Kärkkäinen and Dr. Oleksiy Khriyenko for finding time for regular meetings and fruitful discussions, for inspiring me with nice ideas and ways to implement them, for great support in writing papers and mentoring this work, as well as for friendly attitude.

I am grateful to the reviewers of my dissertation Prof. Francesco Guerra and Assistant Prof. Yuji Tou for their constructive comments. It is an honour to have Docent Kristian Kiili as an opponent and I am very grateful to him.

I thank Prof. Vagan Terziyan and Dr. Olena Kaikova for organizing a student exchange programme between the University of Jyväskylä in Finland and Kharkiv National University of Radio and Electronics in Ukraine. And also for giving me a possibility to complete my Master degree in Finland and start my academic career at the University of Jyväskylä, as well as for great support and creative discussions.

I would like to thank all the co-authors of the papers related to this work: Prof. Pekka Neittaanmäki, Prof. Tommi Kärkkäinen, Prof. Marja Kankaanranta, Dr Oleksiy Khriyenko, Dr. Mirka Saarela, Iida Porokuokka, Renny Lindberg, Ari Tuhkala, Laura Karilainen. I am very grateful to them for being involved in this topic, for discussing inspiring ideas, and developing them together.

I acknowledge the PhD program of the Faculty of Information Technology, COMAS, and Department of Mathematical Information Technology for generously funding my PhD research. I thank these funding parties for financial support in conference participation and INFORTE courses.

I am also grateful to Prof. Heikki Lyyttinen for having an opportunity to work on Adaptive Learning Games , which could help in children development and better learning.

As well, I thank all my colleagues who supported me during my research work, who gave me lots of hints and who shared with me their experiences, and, of course, for the nice and kind atmosphere.

I thank my dear friend Oleg Zolotukhin for great support and help during writing this work, and for providing very useful guidelines.

I would also like to thank Ukraine as my homeland, which had given me an identity, education and vast life experience. I also would like to express my gratitude towards Finland not only as a country where I have been living for last six years, but also as a wonderful place where I got my second Master degree and my job as a researcher.

I am grateful to my beloved boyfriend Andrii Gontarenko for the warmth, motivation, support and belief in my success, as well as for always being there for me when needed.

Finally, my biggest thanks to my loving parents and relatives for giving me vision and motivation to do research, and for giving me strengths to accomplish this wonderful and important step in my life.

Last, but not least, I wish to thank all my friends for their support and friendship.

LIST OF ACRONYMS

AA	Academic Advising
AL	Adaptive Learning
DM	Data Mining
EDM	Educational Data Mining
ILSS	Intelligent Learning Support System
KB	Knowledge Base
LA	Learning Analytics
LE	Learning Environment
LG	Learning Game
LMS	Learning Management System
LO	Learning Object
ML	Machine Learning
OWL	Web Ontology Language
PL	Personalized Learning
PLE	Personal Learning Environment
RDF	Resource Description Framework
RDFS	Resource Description Framework Schema
RS	Recommendation System
SW	Semantic Web
SWRL	Semantic Web Rule Language
URI	Uniform Resource Identifier
W3C	World Wide Web Consortium
XML	Extensible Markup Language

LIST OF FIGURES

LIST OF TABLES

CONTENTS

LIST OF INCLUDED ARTICLES

PI Mariia Gavriushenko, Marja Kankaanranta, Pekka Neittaanmäki. Semantically Enhanced Decision Support for Learning Management Systems. *9th International Conference on Semantic Computing (IEEE ICSC 2015), Anaheim, California, USA*, 2015.

PII Mariia Gavriushenko, Laura Karilainen, Marja Kankaanranta. Adaptive Systems as Enablers of Feedback in English Language Learning Game-based Environments. *2015 IEEE Frontiers in Education Conference (FIE), El Paso, Texas, USA*, 2015.

PIII Mariia Gavriushenko, Oleksiy Khriyenko, Iida Porokuokka. Adaptive Vocabulary Learning Environment for Late Talkers. *8th International Conference on Computer Supported Education (CSEDU 2016), Rome, Italy*, 2016.

PIV Mariia Gavriushenko, Oleksiy Khriyenko, Ari Tuhkala. An Intelligent Learning Support System. *9th International Conference on Computer Supported Education (CSEDU 2017), Porto, Portugal*, 2017.

PV Mariia Gavriushenko, Mirka Saarela, Tommi Kärkkäinen. Supporting Institutional Awareness and Academic Advising using Clustered Study Profiles. *9th International Conference on Computer Supported Education (CSEDU 2017), Porto, Portugal*, 2017.

PVI Mariia Gavriushenko, Renny S.N. Lindberg, Oleksiy Khriyenko. Smart Educational process based on Learning Capabilities. *9th annual International Conference on Education and New Learning Technologies (EDULEARN 2017), Barcelona, Spain*, 2017.

PVII Mariia Gavriushenko, Iida Porokuokka, Oleksiy Khriyenko, Tommi Kärkkäinen. On developing adaptive vocabulary learning game for children with an early language delay. *Manuscript for GSTF Journal on Computing (JoC)*, 2017.

1 INTRODUCTION

It is well known that use of this technology is spreading throughout the world. Web-based learning systems have received increasing attention in recent years. The main advantage of these systems is the ability to learn anytime and anywhere. Learning Management System (LMS)s are good examples of learning with information technology that meets the needs of modern society, taking into account monetary and time costs. Game-based Learning Environment (LE)s are good solutions for knowledge presentation and inspiring learners (Ellis, 2009; Harper et al., 2005; Kankaanranta, 2007; Chen, 2008). However, it is postulated that "one of the main problems with e-LEs is the lack of personalization" (Cristea, 2004; Rumetshofer and Wöß, 2003). Each person has a unique, individual knowledge background, type of memory, learning speed, learning style, motivation for learning, and goals for a future life (for example, preferences in career), as well as different habits, such as when and how to learn. A teacher can adapt the learning course, learning instructions, or learning material according to the majority of learners in the class, but that learning process is not adapted to the personality of each individual learner. The same problem occurs in educational Web applications that were built for a specific group of users and which would be not suitable for others. However, in many cases, the user just works one-on-one with the course or a Web mentor.

The integration of technology in education not only be a new, flexible delivery medium, but also needs to enhance the teaching and learning process by integrating adaptivity and personalization. This goal has driven the development of LE with individualized learning mechanisms became very popular research topic as well as a body of research. Recent advances in technology have contributed to the use of e-learning in education and made personalization possible (Nat et al., 2010). However, it is still a challenging task to develop Personalized Learning (PL) system that include intelligence and adaptivity.

In such LEs, it would be an advantage to monitor the learner's knowledge level and automatically adjust the content for each learner to improve the learning process and make it individual for each user or group of users. Since the Web offers a vast amount of information, it would be also reasonable to help the teacher in the creation of materials most suitable for education and also help in finding the most relevant content and converting it to comprehensive information. It is also necessary to help students with selecting courses and study programs based on their objectives, which will be beneficial in their further career development.

1.1 Research Objective

The objective of this research is to study the challenges of the LEs due to increasing technological progress. From the point of view of a reasonable user with limited time resources, the time spent learning from a personal device must be as effective as possible. In general, efficiency in training systems is understood as the minimum time spent on training: i.e., there must be ways of maximally successful assimilation of information.

Every learner is individual, and has their own learning speed, knowledge background, and preferences. However, learning systems use average learning program determined to be most suitable for everyone in a specific grade. The development and use of LEs with individualized learning mechanisms could assist in online learning and adaptively provide most appropriate learning path for each individual.

Another advantage would be the ability to monitor the learner's knowledge level and automatically adjust the content for each learner to improve the education process and individualize learning individual for each user or group of users.

1.2 Main Contribution

The goal of this study is to address the challenges of learning systems and, subsequently, to propose efficient models of mechanisms for PL. To achieve this goal, the dissertation formulates the following tasks:

- Exploration of Adaptive Learning (AL) system tasks, method, techniques and approaches for building these systems.
- Make the comprehensive theoretical analysis in the area of AL systems, Recommendation System (RS) and feedback systems in LEs.
- Develop an innovative LE model with adaptive tools. These tools would consist of learning feedback system, which would help to create students learning path, provide information about progress, and guide learners during the entire learning period.
- Develop a personalized tool, which would help to detect gaps in the users' learning process and suggest relevant learning material to fill these gaps.

– Develop an AL tool for recommending learning path based on learners' background and preferences.

1.3 The Scope and Research Questions

The research work is focused on adaptive and personalized tools for LE improvement . The scope of the work is limited to designing models of these tools and a general description of the ecosystem of an Intelligent Learning Support System (ILSS).

The main research questions of this thesis are as follows:

1. RQ1: How to make LEs adaptive and what kind of methods, techniques, and approaches could be used for building these systems?
2. RQ2: What kind of feedback is better to use while creating AL systems with a high level of adaptivity?
3. RQ3: How to recommend learning paths for learners, based on the individual's particular skills or preferences?
4. RQ4: How to detect learning gaps and take them into account?
5. RQ5: How to create a tool for adaptive recommendation of the appropriate learning material for slow learners, taking into account different kind of difficulties in the learning process?

1.4 Structure of the Dissertation

The format of this dissertation is a collection of articles. The structure of the work and related papers therein is presented in Figure 1. The original thesis is composed of five chapters. The first is an introduction chapter. Chapter 1 draws a general overview of the current state of learning and growing technology, describes a research gap in this field, and narrows down to the topic of the study. This chapter outlines the main contribution of the conducted research. The purpose of Chapter 2 is to introduce modern LEs and their latest advances. The chapter provides necessary background knowledge about general terms, such as LMS and Learning Game (LG). The chapter concludes by describing the need for adaptation in LEs. Chapter 3 introduces origins and principles of adaptive learning and PL. Chapter 4 focuses on Semantic Web (SW) technologies and Chapter 5 on Learning Analytics (LA), which are presented as instruments for LE adaptation and personalization. Chapter 6 focuses on thesis's research contribution. Chapter 7 concludes the thesis, while Chapter 8 discusses the limitation of the proposed approach and outlines future research.

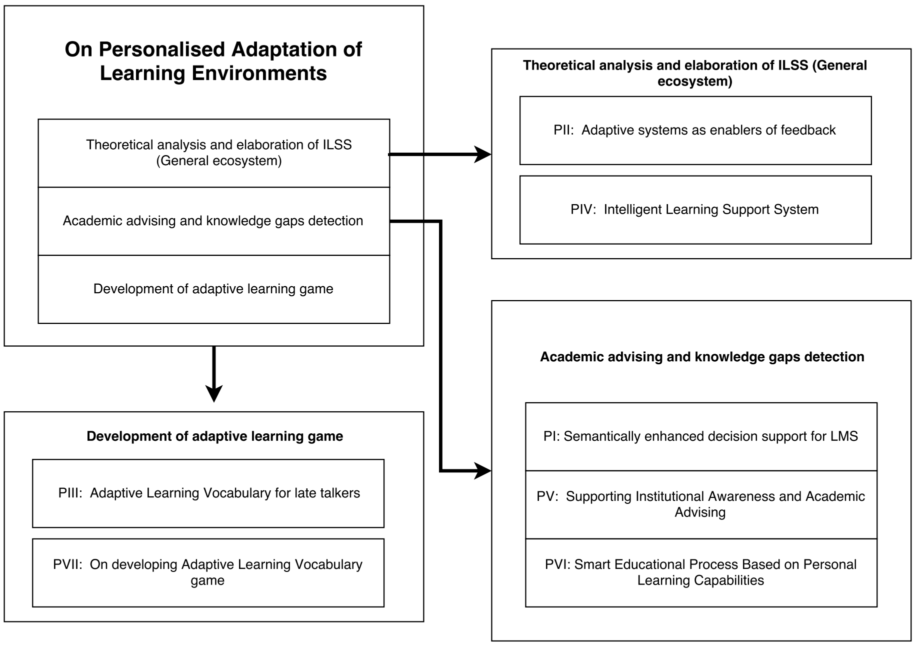


FIGURE 1 The structure of the PhD work

2 LEARNING ENVIRONMENTS

Nowadays exist many types of educational systems: traditional classroom, e-learning, learning management systems, adaptive hypermedia education system, educational game environments, virtual environments, ubiquitous computing environment, etc.

Today's LEs have concentrated more on Industrial Age methods, encouraging learners to remain passive and treating them as if they all have the same learning performance. It means that low-achieving students will stay behind and higher-achieving students will be held back (Reigeluth, 1997; Reigeluth and Garfinkle, 1994).

Integration of information and communication technology into LEs has made important contributions for learning process (Özyurt and Özyurt, 2015). The quality of education is crucial, taking into account that technologies are growing, as is the trust in teachers to deliver a good education that includes advancement of information and communication technology. This should lead to digitally-rich LEs (Watanabe et al., 2017). Another work related to higher education and its quality assurance was discussed in (Terziyan et al., 2015). Authors developed semantic portal as a tool for structural reform of the Ukrainian Educational System. The solution is an ontology-driven portal which is a digital ecosystem for national quality assurance.

In Finland, there are many projects concentrated on innovative and AL. One of the example is USUKO – new generation school "platform for accessing structured, classified and quality-controlled digital learning material for primary and secondary education in Finland" (Kansanaho, 2014). Another project is GraphoGame (Jere-Folotiya, 2014; Richardson and Lyytinen, 2014), a mobile and desktop game that teaches children letter and letter-sound combinations. GraphoGame algorithmically adapts to user skill levels and provides actionable learning analytics for parents and teachers.

In order to make the learning process more customized for each individual there is need to include innovative technology in the LEs.

The evolution of educational technology is as follows:

– The 1990's: Computer-assisted instruction (a PC in every classroom).

– The 2000's: Using LMS (every course has its own Web site).
– The 2011: Massive Online Open Courses (education made widely available).
– Now: AL (education process should be smart). Technology is needed to track learner's progress, analyze behaviour, understand what type of guidance is needed, and provide a relevant learning material. This material may include a personalized learning path and necessary instructions, as well as recommended learning materials and courses tailored to the learners' preferences and abilities.

Lifelong learning is also an important issue that needs to be considered in LEs. It is connected to developing technical and labor markets with knowledge and skills. "The increasing speed of adoption and implementation of new technologies in the workplace and the increasing instability of employment due to the computer driven industrial revolution, it was reasoned that workers would need continuous learning throughout their occupational skills and knowledge to learn new occupational competences" (Attwell, 2007).

Learning quality and learning innovations are relevant topics for a better future of education and training. Without innovated technologies, we cannot participate in the many new changes due to globalization and internet access as well as for the new digital generation (Stracke, 2012).

It is important to analyze learning activities in LEs as well as to make changes in these activities for improvement of learning process. This chapter describes two educational systems, LMSs and LGs, that have increased in popularity in recent years. LMS and game-based LEs are a good examples of learning with information technology adoption. These systems meet the requirements of society, while taking into account monetary and time costs. Also, this chapter presents some challenges in educational systems.

2.1 Learning Management Systems

LMS is a server software that provides a set of interfaces to a database that includes user information, courses, and content. These systems allow educational institutions to manage a huge number of online or mixed courses with commonly used interfaces and resource sets. Offline courses that use LMS as obligatory or advanced tool are often expanded online courses (Kats, 2010).

LMS is a very powerful learning tool that became available with the development of technology. The active development of LMS came in the middle of the 2000s and still increasing in popularity as more people have access to the World Wide Web. Different types of LMS are used by instructors and teachers, from school education to enormous corporations to train and grow their employees.

The survey of LMSs and various plugins used in these systems to improve the educational process is reviewed in the paper **PI**. In addition, a semantically enhanced model of decision support system for LMS is presented in paper **PI**. The proposed system can suggest the kinds of courses necessary to study, in addition to courses that the users have already completed.

2.1.1 Functions and advantages of Learning Management Systems

The main function of LMS is the creation of content, such as generation of course content without the need to master the special skills of a text or HTML editor, or providing communication tools. With an LMS instructors can set up communication channels for student-student or student-instructor interaction; use knowledge testing tools to create tests, complete verious tasks, and persorm administrative functions such as activating or blocking course access, personalizing courses, and managing settings. In addition, LMS allows an administrator to connect additional components (Bailey, 1992; Gilhooly, 2001).

The main advantage of using LMS in the learning process is that is can store an enormous amount of media for information dissemination, communication, and knowledge testing which gathered within a single platform that is publicly available through the Web. In addition, the use of LMS simplifies the process of accessing information for students' courses, and password-based authentication makes it less likely to worry about course material protection.

The LMS system also has some disadvantages. There are a large number of systems aimed primarily at administrative needs, not students' needs. Often the interface implemented by LMS is extremely difficult, and courses represent repositories of different materials that do not meet the learning objectives (Dias and Diniz, 2014).

Modern LMSs are divided into different types according to needs. LMS exists in several types: commercial and free, locally installed, or Software as a Service services (a desktop or Web application), with or without courses, or integrated with other systems.

2.1.2 Learning Management System Trends

By examining a modern LMS, we can distinguish the following trends: gamification, mobile learning, mixed learning, and personalization of the LE.

Gamification is the application of game-design approaches to non-gaming applications. Gaming makes the learning process more fun and rewarding, since instructors and students get active feedback, maintain the spirit of competition between students, and determine the complexity of the material depending on the skills of students. At the same time, gamification has its disadvantages, expressed in the competitive pressure on students, the difference between course material and supply in the market, and the reduction of attention (use of gaming accelerates training, and students will expect the same speed from other educational activities).

Mobile learning includes the use of mobile gadgets for learning. The main advantage is that it allows studying wherever and whenever. Examples of this type of training are the dissemination of course materials through the network, interaction with the audience during the lesson and the instant receipt of a teacher's feed from the course during an online lecture. The advantages of this type of training are the ability for the student to choose the time and place of study, the variety of content, and the opportunity to work with teachers from other countries. This method's disadvantages include limiting the volume for the devices' battery, as well as the inability to control what the students themselves are engaged in because they are tempted to spend time on the gadget, but not for training purposes.

Combined learning blends student and instructor meetings with distance learning. This approach allows the combination of different models, but all allow you to increase the number of groups, reduce the cost of student training, and allow each student to work at his own pace. However, there are also some disadvantages: the lack of motivation, the need for basic knowledge of technology, and the problem of finding information in inappropriate sources.

A Personal Learning Environment (PLE) is the use of past student learning experience to individualize the learning process. As part of the PLE approach, a model where the student is the central element of learning, rather than the course. It became available with the development of Web 2.0 technologies. The most common PLE template is a Web-based application that are similar to social networking. The PLE approach helps to set certain student learning goals that are accessible not only to the instructor but also to the student. A special feature is a need for a high level of self-discipline on the part of the student (Ndongfack, 2016).

2.2 Learning Games

Playing games is an important part of our social and mental development. Recently, educational games have become more popular (Kankaanranta et al., 2017). There are many games specifically designed as educational games, as well as a number of entertainment games that have been successfully used for educational purposes (Comunale, 2017).

Gamification has a huge impact on teaching children. Taking into account the specifics of the modern impact of information technology on the educational system (for example, mobile phone development is growing due to innovative technologies (Watanabe et al., 2012)), the structure of gaming technologies is changing. The various aspects of the influence of learning computer games are significant, such as influence on the learners' development, their mental and psychological health, thinking, and memory.

LGs are software that trains and educates a person in the game mode, and can be used for both training and entertainment. The category of the training game includes genres such as quest, arcade, 3D-shooter, simulator, and an interactive course on any subject. The program divides the teaching material into parts, and regulates the sequence of its study. The assimilation of the material is checked by the test offered at the end of each stage of the training (Ghazal and

Singh, 2016).

Digital games are recognized as high potential educational tools which could provide learners with new and increasing learning opportunities (Bellotti et al., 2011). Digital games utilized in education can be divided between mainstream games which are exclusively created for fun and LGs, which are obviously created for educational purposes (Kirriemuir and McFarlane, 2004). LGs also refer to educational games and serious games (Breuer and Bente, 2010). They can create a new learning culture that corresponds better to learners' interests and habits (Prensky, 2003).

In the LG, it is possible to acquire and consolidate knowledge and skills through activities, according to specified rules. There are two allocated components: teaching and playing. In the lesson, one of the components can predominate (i.e., it can be a game for training and learning while playing). If the training component prevails, then the game provides ample opportunities, associated with the perception of knowledge. In the case of predominance, the game component can be used as a tool for visibility and enhancement of the motivation to learn.

LGs present an important possibility for education improvement and also in the lifelong learning perspective (Bellotti et al., 2011).

The most important functions of the LG are (Bellotti et al., 2011):

1. The training allows to solve educational and training tasks aimed at the assimilation of certain program material and rules that should be followed by the players. LGs are also important for the moral and aesthetic education of children.
2. The entertaining function helps to increase emotional and positive tone, development of motor activity, nourishes the child's mind with unexpected and vivid impressions, and creates a fertile ground for establishing between the adult and the child.
3. The communicative function consists of the developmental needs to exchange knowledge and skills with peers in the process of games, communicate with them, and establish a friendly relationship, to display speech activity.
4. The educational function helps to identify the individual characteristics of children, and it helps to eliminate undesirable manifestations in the character of their pupils.
5. The developmental function is connected with the development of the child.
6. The relaxation function restores the child's physical and spiritual strength.
7. The psychological function develops the creative abilities of children.

There are identified categories which address decisive educational aspects that designers of LGs should take into account during the game design (Kiili, 2010):

1. Integration patterns, which present solutions that harmoniously combine game elements and learning objectives in pedagogically meaningful ways.

2. Cognition patterns, which motivate reflective and meta-cognitive processes in learners and motivate learners to process relevant content.
3. Presentation patterns, which goal is to make sure that the learner's content processing is effective.
4. Social interaction and teaching patters, which aim at trigger reflective and meta-cognitive processes through social activities and socially created game elements.
5. Engagement patterns, which package all gaming experience into meaningful and motivating package.

One of the key focuses in developing LG is design items and choices related to the content presentation (Kiili, 2007):

1. Interface: Graphics which attract the learner, including engagement in play, but also should be used appropriately so the learner can concentrate on the learning process.
2. Method/level of interaction: Interaction modalities such as question-answering or exploratory strategy, which should be determined by the game's inspiring pedagogical principles.
3. Feedback and reward: A a consequence of the learners' actions, which could be a simple yes or no answer or an explanation of the answer and an adaptation of learning process. More detailed information of feedback in LG can be found in paper **PII**.
4. Types of evaluation of performance and level of meta-cognitive support provided: The performance of the system and its ability to provide learners with a proper evaluation of their performance.

Game development for children, especially young ones, is still an emerging market; however, digital games comprise a fast-expanding field (Kankaanranta et al., 2017). In paper **PIII** and **PVII**, we described the vocabulary learning game for late talkers. Paper **PIII** concentrated on the architecture creation for AL and three dimensions of similarities for the words which could influence learner's speed of learning. Paper **PVII** concentrated more on the development of adaptive vocabulary learning based on semantic similarity.

The learning system developed like a story, in which the animated character saves animals which he or she has to study. The filtering, learning, and testing process is described. While these processes are passing, child is earning coins, so he or she can spend them after for getting new cloth for animated character, new colors, etc. The filtering, learning and testing trials are short, so the child will not get bored. The learning process of the game is presented in Figure 2. The learner is learning the word "leopard", and the system shows this word among the distraction words, which are not similar to leopard. Mostly cold-blooded animals are shown as distractors in this case. The game-based learning system, which takes semantic similarities and other similarities into account, also adaptively changes the distance between the words, according to the child's progress and abilities.

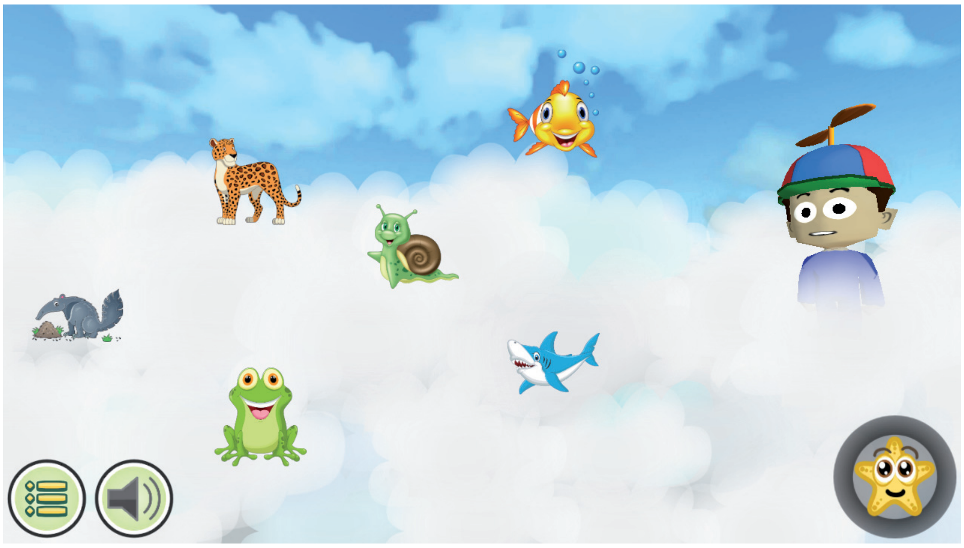


FIGURE 2 Learning process of adaptive vocabulary learning game

3 TOWARD ADAPTIVE LEARNING

This chapter describes adaptive and personalized systems, main concepts, tasks, features, core elements, layers, and techniques.

3.1 Adaptive and personalized learning systems

Every system requires the collection of some elements that are in relationships and connection with each other, and these elements form a certain integrity. The computer system is a functional block which has input, output and logic inside the block. The adaptive system will change its own functionality according to input, which also will influence on the output. According to McGraw-Hill Dictionary[1], an adaptive system is a set of items which connects to each other and interacts to each other, and one of the processes of the system is responsible for the adaptation by the correlation between structure, function or behavior and its environment, to improve effectiveness for achieving goals. The importance of these systems can be interpreted as a result of the impulse given by the broad spread of the World Wide Web (García-Barrios et al., 2005). It is seen, that there is a real need for such systems in different situations of modern life. Adaptive systems can adapt to different contexts, such as a person, devices, software, or various environments. Particularly, the high significance of adaptive systems can be identified in the field of e-commerce and adaptive e-learning (Kobsa, 2001).

AL systems are technological systems that are capable of adapting and learning depending on received input signals (Sonwalkar, 2005). The main tasks of adaptive systems are: to allow organizing content, offer learners different contexts and perspectives, recognize the way a learner likes to learn by completing diagnostic assessment of the learning preference, use assessment results for continuous intelligent feedback that motivates and provides guidance to overcome concept deficiencies, and help to maximize learning performance.

Personalized systems present a specific type of general adaptation systems

[1] McGraw-Hill Dictionary of Scientific and Technical Terms

(García-Barrios et al., 2005). In general, personalization means adaptation for a specific user. Personalized systems have had a great success in areas as e-learning and targeted advertising; and e-learning. Personalization in the field of education means an automated way of teaching that adapts to different qualities of learners, and it take place when learning becomes personal in learner's mind. Personalization can be described automatic learning paths structured to meet the needs of the learner.

There are two types of adaptation for personalization in learning systems (Santally and Senteni, 2005): when the system allows the user to change certain system parameters and personalize its behavior (person-driven personalized system), and when the system adapts to the users automatically according to the users' needs and actions (system-driven personalization). In the first case, the learner is responsible of the learning process and of the tools which they use. In the second case system is more intelligent, it takes each decision a learner makes and adapts the learner's learning pathway, both within lessons and between lessons, thus producing many PL paths, each specifically designed for a learner's unique needs in real time. This is especially relevant when a person cannot formulate personalization alone. A self-adaptive system is a system that "sense" the user. By sensing we mean input of user's information (knowing the information about the user). This information could be direct or indirect. Direct information (such as goal-driven information) means that user can provide some information about his preferences (like preferences for future specialization). Indirect means that the system can extract the information from the Knowledge Base (KB) or according to the user's actions.

Personalization could also be based on extracted data from the user's profile, such as knowledge background or set of preferences provided by the user, also according to the analysis of user's actions such as analysis of a user's responses or mistakes, or could be based on both, depending on the learning objectives that were initially set in the system. For example, if the learner's goal is to receive knowledge in some specific field for getting the desired specialization, then the system should adapt study content according to the learner's knowledge background and preferences. The main function of the system will be selection of content. If the system's goal is to teach the learner in an optimal way, then system should take into account not only extracted knowledge background of the learner, but also his personal ability for learning new material. Thus, we must minimize the time and maximize the quality of the learning content. System should improve the logic of the learning concepts' presentation, taking into account personal specifics of the user via manipulation of the complexity level of the system.

3.2 Tasks, features and core elements of Adaptive Learning Systems

Adaptation in learning has multiple aspects:

- Adaptation to the current needs of particular learner.
- Adaptation to a particular learner's state.
- Adaptation to a chosen specific field for studying.
- Adaptation to a specific task.

According to (Sonwalkar, 2005), the goal of an AL system is to personalize instructions in order to improve the learner's performance, identify the strengths and weaknesses of the learner in understanding, provide relevant learning material until defined goals are achieved, and enhancing the individual learning process, taking into account the speed of the learner, accuracy, and quality of learning (Oxman et al., 2014).

The main tasks of adaptive systems are as follows (Sonwalkar, 2005):

- Allowing organization of content in various ways, offering learners different contexts and perspectives.
- Identifying the way a learner prefers to learn by organizing assessment of the learning preference.
- Using assessment results to provide intelligent feedback that motivates and helps in maximizing learning performance.

Adaptive learning systems have a range of features and functions (Venable, 2011):

1. Pre-test: The system assesses the current knowledge and skills of the learner. These assessments gather information about individual learner characteristics, including prior knowledge.
2. Pacing and control: The ability of the learner to control the speed of learning at which the content is delivered and also to choice within the system.
3. Feedback and assessment: system continuously assesses learner's progress, provides feedback about learner's work: feedback can include information about whether the answer was correct or not, also suggest information on resources for opportunities for additional practice.
4. Progress tracking and reports: The system tracks individual progress and allows learners to return to the point where they left off, breaking the work into multiple sessions. The system creates periodic reports for instructors for future guidance of learners based on their performance.
5. Motivation and reward: Adding gamification to the system in a way of rewarding learners with points or badges.

Core elements of the adaptive system (Oxman et al., 2014):

1. A content model that refers to the way of structuring the topics or content domains with detailed learning outcomes and a defined tasks that must be learned. The sequence of content could be changed according of the learner's performance.
2. A learner model, which refers to estimation of learner's ability level on different topics, tracking their performance and tracking existing KB about mastered topics.
3. An instructional model, which determines how a system selects specific content for a specific student at a specific time. In other words, it puts together the information from the learner model and content model to generate the learning feedback or activity that will be most likely to advance the student's learning.

3.3 Types, layers and techniques of adaptive systems

One possible division by types of AL systems is related to types of learner model (Weber and Brusilovsky, 2001): stereotype, overlay, case-based and Bayesian learner models.

1. Stereotype user models are basic learner models that arrange learners by capacities utilizing data which was accumulated ahead of time. For instance, it can be the favored learning style of the student, a suitable learning way, or inspiration and so on.
2. Overlay user models as shown by a model of the thoughts and ideas of an area to be educated. These models are often used in adaptable learning systems.
3. Case-based user models give feedback to the users about right or wrong answers or solutions. These models have been utilized as a part of intelligent tutoring systems.
4. Bayesian user models, particularly Bayesian knowledge tracing, has been used successfully in several cognitive tutors and in many intelligent tutoring systems.

According to (Oxman et al., 2014) the division of AL systems are as follows:

1. Simpler AL systems: are rule-based, which means a series of if-then statements. These systems are content-oriented and easier to understand in terms of functionality, but are limited in their ability to adapt to individual learners' abilities or preferences. Rule-based systems are mostly built using a series of if-then functions. These systems employ a straightforward branching architecture. A learner is asked a question; if they answer correctly, they move on to the next selected question; if they answer correctly,

they move on to the next selected activity, but if not, they receive additional content.

2. Algorithm-based systems are more complex than rule-based systems. They use advanced mathematical formulas and Machine Learning (ML) concepts to adapt to individual learners with greater specificity. Mathematical functions are used to analyze learners' performance and content. At their most sophisticated, these AL systems involve ML capabilities where the system learns more about the student and the content as the lessons progress. Such systems use LA to deal with big data, and employ algorithms to predict the probability of a particular learner's success, based on specific content. These systems may classify learners and learning content, and they are more complex one according to the number of variables they are taking into account. To make the adaptive process more powerful, cloud computing technologies and techniques for analysis of big data have been used. For better interpretation of spoken or written learner input, the system could use natural language processing.

There has been also described five layers of adaptivity (Oxman et al., 2014):

1. Learners complete quizzes after studying the material, and the system provides feedback on whether the learner answered correctly.
2. Learners complete quizzes after studying the material, and the system provides feedback on whether the learner answers correctly. The system also provides some explanation to their incorrect answer and offers study material to support learning or review of the material.
3. Learners complete quizzes after studying the material, and the system provides feedback on whether the learner answers correctly. In the next quiz the system will add the questions that have been answered incorrectly in the previous section to see if the learner mastered them. The learners can track their progress.
4. Learners learn some material and then take quizzes. Based on the learners' results, the system provides an individualized learning plan to fill in gaps. After passing this plan, learners take another quiz that ensure they make progress.
5. A learner is given a math or science problem and is asked to fill empty slots with the steps towards the solution. Feedback could be given immediately after each step or after whole qiuz. The feedback is based on analysis of the answers the student entered into the slot, compared to the most effective methods for solving such a problem. After that, the system selects the next question based on the learners' readiness.

More detailed information about layers and the feedback types connected to them are presented in **PII**.

Technologies used in AL systems include the following: Semantic Web and rule-based reasoning, Bayesian networks, adaptive hypermedia, service-oriented architecture, agent-based systems, ML and data mining, user and group modeling, empirical modeling, context modeling and prediction, quality of service management, and decision support. In this Ph.D. work, we used mostly Semantic Web Technologies and Data Mining techniques, which are useful in knowledge presentation and knowledge extraction (creation) as well as in the analysis of learners' behavior and in decision support.

4 SEMANTIC WEB

SW technologies aim at adding semantic information to Web contents for creation an environment where software agents can communicate with each other and do tasks more productive. SW was intended by Tim Berners-Lee; it is a group of technologies and methods that allow machines to understand the meaning of information on the Web and integration of information in an intelligent way (Castellanos-Nieves et al., 2011; Bouquet and Molinari, 2012; Berners-Lee et al., 2001). SW provides environments where software agents can navigate through Web (Bittencourt et al., 2008).

SW technologies includes Resource Description Framework (RDF), Resource Description Framework Schema (RDFS), Web Ontology Language (OWL), Uniform Resource Identifier (URI), Extensible Markup Language (XML), and SPARQL (Ahmed et al., 2013). SW allows using more automated functions on the Web such as reasoning, autonomous agents, information and service discovery (Berners-Lee et al., 2001).

This chapter presents a brief review of the main technologies of SW and their use in LEs, more precisely using RDF, ontologies, and semantic annotations to empower learning systems.

4.1 Semantic Web in learning environments

SW technologies have been used in LEs in various ways. When using ontologies, it is possible to create a semantic model that consists of concepts, their properties and relation, as well as axioms related to previous elements (Castellanos-Nieves et al., 2011). Approaches which could enrich, categorize and retrieve Learning Object (LO)s attracted attention recently (Halimi et al., 2014). SW technologies are viewed today as those that can resolve the problem of interoperability and integration within a heterogeneous world.

SW technologies, such as ontologies and semantic annotations, have been used in LEs for designing and classification of learning material (Devedžic, 2006;

Guangzuo et al., 2004; Lee et al., 2008; Lytras et al., 2003; Wang and Hsu, 2006). Ontology-based personalization plays a leading role in building smart e-learning ecosystems (Ouf et al., 2017). If LOs include ontology-based properties and hierarchical semantic associations, then the learning system will have the ability to provide adaptable and intelligent learning to learners. SW technology can provide useful relevant information for searching and sequencing learning materials in learning systems. Learning systems enriched with SW could provide adaptable and personalized LO to learners (Ahmed et al., 2013).

With SW, it is possible to make learning smarter since data can become knowledge, which in turn can be used as content for learners at any time and in any location. SW technologies are great way to empower learning systems by providing educational resources dynamically, find-able and adaptive to users (Kohlhase and Kohlhase, 2008; Markellou et al., 2005; Berners-Lee, 2005).

4.2 Resource Description Framework

A Recourse Description Framework (RDF) is an abstract model. According to this model, knowledge is represented by a triplet in the form of a relationship: subject – predicate – object. Thus, in the RDF model, the nodes of the graph are subjects and objects. The edge always has a direction to the object. A predicate is also often called a property. Each object and subject can be specified both as a URI and as a symbol and, for some tasks, could be empty. The predicate must have a unique URI (URL). It has its own description, identified by the URI for a clear understanding (for example, systems of semantic analysis) (Carapina et al., 2013).

The subject, predicate and object are the names, which could be global, referring to the same entity in all RDF documents where it is used, or local, where the entity referred to by this name cannot be referenced directly from outside the RDF document.

Areas of use:

1. Combining data from various sources without resorting to the creation of specialized programs.
2. Need to give access to your data for others.
3. Need to decentralize data so that they are not "owned" by anyone alone.
4. Need to do something special with large amounts of data, such as to enter, retrieve, view, analyze, or perform a search, etc.

Using an RDF model, it is possible to get:

1. Logical conclusion of new facts.
2. Ensuring Semantic Search.
3. Flexibility of the data model.
4. Extreme ease of data exchange between systems.

The RDF model provides formal descriptions, which means that the search agent can search for the facts and knowledge. In other words, RDF is a universal way of decomposing any knowledge into small pieces. It sets certain rules regarding semantics, i.e. sense of these pieces. The idea is that, in one simple way, any fact in this structured form could be processed by computer programs.

Using RDF it is possible to describe documents, individual fragments of knowledge within a document, or objects in the real world, such as a specific living person.

RDF is suitable for working with distributed knowledge because applications can collect RDF files placed on the Web by different people, but it is easy to recognize elements that were not in any of its parts from the collected document. In RDF, there are two processes that make this happen: first, documents that use common languages are combined; and second, it is possible to use any languages in each of the documents. This flexibility is one of the distinguishing features of RDF (Stojanovic et al., 2001).

4.3 Ontology

Ontology is the core element of SW. An ontology defines relations between terms. The main uses of ontologies are to enhance the accuracy of searches and reuse of knowledge representation. Ontologies could be general, top-level, or domain-specific (Gruber, 1995).

In general, the structure of an ontology is a set of elements of four categories:

1. Concepts.
2. Relations.
3. The axioms.
4. Individuals.

Concepts are considered as a conceptualization of the class of all representatives of a certain entity or phenomenon (for example, Animal or Feeling). Classes (or concepts) are general categories that can be hierarchically ordered. Each class describes a group of individual entities that are combined based on the availability of common properties.

Concepts can be associated with various kinds of relationships (for example, Length or Location) that link classes together and describe them. The most common type of relationship used in all ontologies is the relation of categorization: that is, referring to a particular category.

Axioms define the conditions for correlating categories and relationships, they express obvious assertions linking concepts and relationships. An axiom can be understood as an assertion introduced into an ontology in a complete form, from which other statements can be derived. They allow expression of information that cannot be reflected in the ontology by constructing a hierarchy of concepts and establishing different relationships between concepts.

The following statement is an axiom: "If X is mortal, then X will die someday". Axioms allow further inferences in the framework of ontology. They can provide researchers with information about rules that automatically add information. Axioms can also be constraints imposed on any relationship that makes it possible to deduce inferences.

Conceptual restrictions indicate what type of concepts a given relation can express (for example, the "Color property" can only be expressed by the concepts of the "Color category"). An example of numerical limitations is the assertion that the number of biological parents for a person is 2. The number and degree of axioms' details usually depend on the type of ontology.

Along with the specified ontology elements, ontology also includes so-called "instances" that may also be called:

– Specific specimens.
– Instances.
– Individual copies.

Instances are individual representatives of a class of entities or phenomena, that is, specific elements of a category (for example, an instance of the class "Man" will be "Queen Victoria"). The components of the ontology are subject to a kind of hierarchy. At the lower level of this hierarchical ladder are specimens, or concrete individuals; above, there are concepts or categories. On a higher level, the relations between these concepts are located, and the step of rules or axioms is the generalizing and connecting one.

By using ontologies, it is possible to enhance the semantic representation of knowledge for association formal descriptions of LOs, to make reasoning, and to search relevant learning materials, compose new learning path from existing resources (Ouf et al., 2017). The benefits of using ontologies are in their re-usability and sharability. Also, ontologies allow access to preferred content (Khozooyi et al., 2012).

The main advantages of ontologies are a better understanding of information in the document, capturing the contextual relationship among various components within the document, categorization, and contextual search. These options improve the quality of the output in many ways (Khozooyi et al., 2012; Chakraborty and Mittal, 2010).

Ontological presentation of the learners, courses and specializations are used in PI. Semantic Web Rule Language (SWRL) it was possible to create rules for new knowledge about specializations and what kind of courses should be included in order to get this specific specialization. In PVII and PIII, the ontology of animals was used for counting semantic similarities between concepts (in our particular case, words). An example of ontological presentation of ILSS is presented in **PVI**.

Ontology could be beneficial in a learning ecosystem, because it could adapt to achieve many tasks, including representing the LOs as taxonomies, annotating, and indexing them for collaborative use, accessibility, and other functions (Capuano et al., 2009).

4.4 Semantic Annotations

An annotation is a form of attaching information to an existing resource. Used in the text, semantic annotation annotates all mentions of instances relating to concepts in the ontology (without modifying the ontology). Semantic annotations are metadata associated with particular information concepts, which are defined by ontology and could be used for semantic enrichment (Pahl and Holohan, 2009).

Semantic annotation attaches metadata to the documents, pointing to concepts and properties in an ontology. Information is typically exported as text annotated with links to the ontology.

There are two ways to create semantic annotations:

1. Manually: ontology based annotation tool.
2. Automatically:
   - Gazetteer/rule/pattern based.
   - Classifier (ML) based.
   - Combinations thereof.

Annotating LOs is a fundamental task, which helps in accessing, sharing, and reusing of learning resources. Annotations allow transforming texts into a structured, semantically enabled content (Bouquet and Molinari, 2012).

Semantic annotations allow robust and scalable handling of material, flexible solution, and reuse of material in different content.

In paper **PVI** we used semantic annotation as a key instrument for the creation of ILSS. It will help to annotate knowledge skills, courses, and learning programs for better identification of relevant learning material, learning paths, and other uses.

5 LEARNING ANALYTICS AS AN INSTRUMENT FOR ANALYZING LEARNERS' DATA

Recently, large amounts of data that cannot be processed by databases alone has become common. This type of dataset is processed by data mining; in the educational field, Educational Data Mining (EDM) and the related field LA (Heiner et al., 2007) are significantly involved. These are very fast growing technologies and they are making a big research interest. Usually, they use ML algorithms, data-mining algorithm, statistics and also semantic technologies. These algorithms help solve educational questions, provide tools and technologies, and develop platforms for better learning processes and experiences as well as preparing learners for success and helping teachers during the teaching process.

This chapter presents a brief origin of the term LA and a description of how it helps to analyze learners' data.

5.1 Learning Analytics

Educational data is growing, and there is a need to understand and leverage this data. LA recognizes the meaning and actionable patterns from a large pool of educational datasets for future decision-making about learning and teaching (Pardo and Teasley, 2014; Gray, 2014; Siemens, 2012).

According to the 1st International Conference on LA and Knowledge (LAK11)[1] held between 27th February 2011 to 1st March 2011, LA is the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs.

The dissemination of LEs in all sectors of education produces huge amounts of tracking data. E-LEs store user data automatically, but maintenance of the data for learning and teaching is still very limited. The importance of LA is grow-

1 https://tekri.athabascau.ca/analytics/about

ing because educational datasets offer new opportunities for learning evaluation, feedback and support, recommendations and learning path creation, early warning systems, and automatization of the learning process.

LA is an independent, growing discipline, which is closely related to EDM (Saarela, 2017). The purpose of LA is to collect, analyze and report data about the learners and their actions for future understanding and changing of learning process in an optimal way. LA studies are increasingly leveraging big data.

The main steps of LA are as follows (Sudarshan et al., 2012):

1. Data collection, which is based on learnersâ actions, answers, and solved tasks.
2. Analysis, which provides predictions of learners' learning activity or detection of learning gaps. This analysis could be presented with visualizations, charts, or tables.
3. Student learning, which, based on the analysis (outcomes), develops content pertaining to each learner.

There are three main advantages in using LA:

1. Learners receive support for the specific learning process and receive PL recommendations for suitable learning resources, learning paths, or peer learners who have similar learning paths.
2. Teachers can monitor systems that inform them about knowledge gaps of particular learners and thus help them to focus their attention. This monitoring learners for better curriculum design and on-the-fly adaptations.
3. LE can monitor the performance of learners regarding graduation rate and drop-outs since it is possible to evaluate courses and improve outcomes of the courses.

The main benefit of LA, compared to general educational research is that LA handles big data, makes sense of it and makes use of it (Ferguson, 2012a).

The main objectives of LA are reflection and prediction. Reflection means critical self-evaluation of a learner's data as indicated by their own data sets in order to obtain self-knowledge. It means self-observation and reacting to one's own performance log data. Apart from support for reflective practice, LA can also be used to predict and model learner activities (Siemens, 2012). This can lead to earlier intervention (e.g., to prevent drop-out), or to adaptive services and curricula, along with computer interaction systems that support this process.

Taking into account these observations, LA is new and interesting area of research with huge interest within the academic community and beyond, and beyond. LA can tailor educational opportunities to each learner's level of need and ability (Ferguson, 2012b). LA provides good possibilities for helping learners and teachers in the educational process. LA has already developed a wide range of tools and methods that express the existing potential of LA.

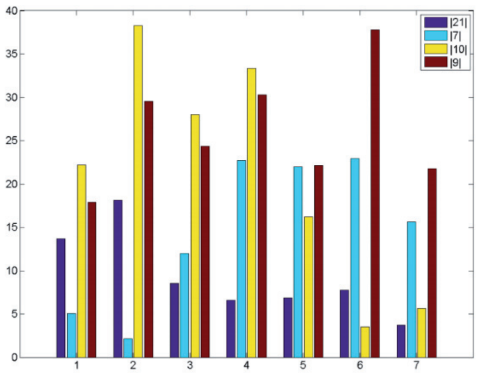


FIGURE 3 Credit accumulation for autumn semester year 2012

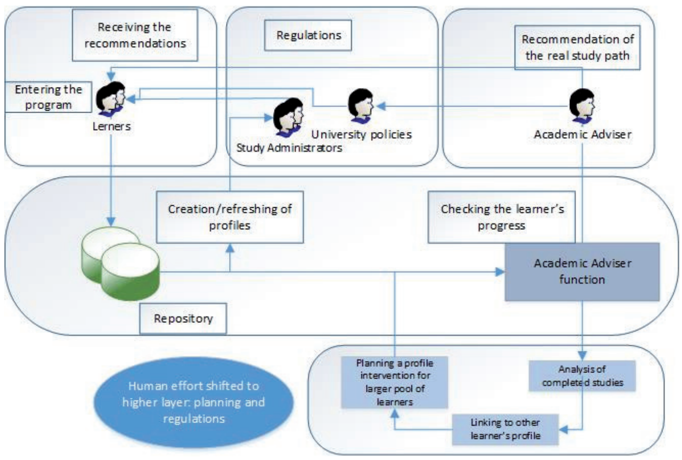


FIGURE 4 Automated process of AA

5.2 Methods and techniques

There are several methods of LA: prediction, clustering, relationship mining, and discovery with models. There are also subjects of LA: assessment of the student's learning behavior, course recommendation and adaptation to learners' behavior, evaluation of learning material, feedback to both teacher and students in e-learning courses, and detection of atypical students' learning behavior (Castro et al., 2007; Chatti et al., 2012).

LA became beneficial in information retrieval technologies like EDM in the form of classification, clustering, and association rule mining (Romero and Ventura, 2010), ML, or classical statistical analysis techniques, social network analysis (Shum and Ferguson, 2012), and Natural Language Processing. According to (Rogers, 2015) knowledge discovery from databases is increasingly important technique of LA. According to (Siemens, 2013), Data Mining (DM) is one of the most sophisticated quantitative methods in LA. Researchers use large databases for analysis, such as analyzing discourses, predictive models, social networks, intelligent curriculum, or sentiments. Using these technologies, LA can contribute tailored information support systems to the stakeholders and report on demand. According to research and practice and taking into account emerging evidence, researchers suggests that LA can be more personal, convenient, and engaging, and may impact on student retention.

Clustering is one of the key techniques in the DM category of LA methods. In **PV**, the process of automated AA was presented. This process is based on LA, specifically using robust clustering. The goal of this paper was to assist academic advisers by recommending suitable courses for students based on passed courses of (possibly more advanced) students with similar study paths. With clustering, it was possible to identify general profiles of similar students. Figure 3 shows the profiles for autumn semester 2012, and the size of the cluster is shown in the top-right corner. The profiles are sorted in the ascending order with respect to the total number of credits. in Figure 4 shows the automated process of AA, which help to plan when and how to provide the courses, especially the compulsory ones, as well as to plan a profile intervention adaptively for a larger pool of students, which will reduce the human effort of academic advising.

6 RESEARCH CONTRIBUTION

This chapter provides an overview of the seven included publications that contain the contribution. It explains the theoretical analysis of ILSS (Section 6.1), academic advising and knowledge gap detection (Section 6.2) and development of an AL game as a use-case (Section 6.3). In the final section of this chapter, the main results are tabularly summarized (Section 6.4)

6.1 Theoretical analysis of Intelligent Learning Support System

This section describes the technologies, features, and levels of AL systems, as well as the ecosystem of ILSS and its main modules.

6.1.1 Article PII: Adaptive Systems as Enablers of Feedback in English Language Learning Game-based Environments

This article was published in the proceedings of the Frontiers in Education Conference (FIE), 2015 IEEE, pages 1–8.

Objectives

The aim of this article is to determine the importance of AL environments to enhance more personalized education as well as to determine the main categories, layers, and types of AL systems. In addition, this article examines the significance of feedback in these systems and makes an evaluation criterion for AL environments using different feedback types and features.

Findings and Contributions

The paper provides a short review of existing adaptive game-based learning systems, specifically concentrated on English language learning.

This work introduces an evaluation framework for the analysis of game-based LEs by the main criteria used in AL games. The analysis is focused on various LEs, specifically applications and games for English learning. The analysis results in an evaluation of 23 game-based LEs and applications that were developed for the purposes of English learning. The relationship between the feedback types and areas of English learning is presented.

According to this research the most common feedback types are informative feedback and point-based feedback. Diagnostic feedback appears only in high-level AL systems. Also, some similarities in feedback types were found for a specific English learning area. According to the results, the more complex feedback types the system has, the more adaptive it will be. Therefore, it is very important to know, which types of feedback to use while creating a learning system.

In this work, guidelines for the developers of various applications and games for English language learning were created. The users of the applications and games will also find information about which learning system to select for better guidance and support.

Author's contributions

The author of this thesis is the main and corresponding author of this publication. She searched for game-based LEs, analyzed them, carried out the data analysis, produced the results, and wrote the majority of the paper. The literature review on English learning was mostly written by the third author, who also helped with the analysis of game-based learning. The third author revised of the paper and helped present the results. The main author also presented the paper at the 2015 IEEE Frontiers in Education Conference in El Paso; Texas, USA.

6.1.2 Article PVI: An Intelligent Learning Support System

This article was published in the proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017), pages 217–225.

Objectives

The focus of this study is to take into account changing labor markets, personal needs, and skills of the learner. The goal of this paper is to propose an ILSS which could deal with learners' knowledge gaps and support the learning process throughout lifelong learning.

Findings and Contributions

The paper addresses important issues in the field of Technology Enhanced Learning. Guiding learners during their whole life-cycle with recommendations for learning materials to fill their learning gaps is described and proposed in this work. The domain problems, motivation of the creation of intelligent learning system is described. This paper presents an ILSS: specifically, a meta-system with main modules that include a Courses and Learning Material Annotation module, an Annotated Learning Content module, a Personal Learner Profile module, and a Personal Study Planner module. This work explains an electronic education system, learning process problems, needs of innovative learning platforms, and the benefit of the system for teachers, students, and government employers. The paper also outlines the research questions. The paper proposes an overview of an intelligent support system for learning that aims to cover a person's complete path of studies from elementary schooling up to professional and career development. The main idea involves annotating the whole learning content, courses, materials, and other associated content. SW, in the form of ontology, is used for annotation. The proposed model of the ILSS could fulfill the needs of modern society, where learning and teaching are not bounded to single institution.

Author's contributions

The author of this thesis is the main and corresponding author of this publication. She conducted the literature review on research related to intelligent systems, wrote the main text of the paper, described the model of of the ecosystem of ILSS, and described the main method of the system: specifically, semantic annotation. The second author also worked on the functionality of different modules of the system. Moreover, the main author presented this paper at the 9th International Conference on Computer Supported Education conference (CSEDU), Porto, Portugal, 2017.

6.2 Academic advising and knowledge gap detection

This section presents work related to the recommendation of study path based on learners' preferences, knowledge background and similarityÐ$\pm andknowledgegaps$.

6.2.1 Article PI: Semantically Enhanced Decision Support for Learning Management Systems

This article was published in the proceedings of the International Conference on Semantic Computing (ICSC), 2015 IEEE, pages 298–305.

Objectives

The focus of this study is the area of LMS and various plugins used in these systems to improve the educational process, as well as to determine the current status of these systems in the industrial and research communities. The idea of this study is on developing a model of intelligent tool for the design of LMS based on the study of semantically enhances decision support.

Findings and Contributions

This paper provides a broad discussion of integrating semantic technologies with LMSs. A variety of technologies are discussed, mostly drawn from SW side. As well, some background on LMS and RSs is included. A model of RSs was suggested for LMS with decision support. The proposed model functions as a study adviser, as it can suggest the kinds of courses necessary to study in addition to courses that the users have already completed.

The developed method of decision support takes into account the interests of users while they select the desired specialization and provide some input in choosing courses. It demonstrates an advanced capability of LMSs: implementations of the decision support method for users. The method was also designed for user convenience. It supports users while they choose the most suitable courses as well as increases the popularity of LMSs. In the field of education, the designed model allows making recommendations based on user data extracted from profiles and user's preferences.

Author's contributions

The author of this thesis is the main and corresponding author of this publication. She searched and conducted the literature review on research related to LMS and different improvements, and developed a model of an intelligent tool for LMS design in a form of RS. Moreover, the author presented this paper virtually at the IEEE International Conference on Semantic Computing (ICSC), 2015.

6.2.2 Article PV: Supporting Institutional Awareness and Academic Advising Using Clustered Study Profiles

This article was published in the full paper proceedings of the 9th International Conference on Computer Supported Education (CSEDU 2017). Moreover, an extended version of this paper has been selected and submitted to the Communications in Computer and Information Science series publisher by Springer, pages 35-46.

Objectives

AA is an iterative, collaborative process between the student, academic adviser, and academic institution. The function is to ensure that the students complete the required studies to graduate. The purpose of this article is to support students with guidance in their study path by proposing an automated academic advising system utilizing prototype based clustering.

Findings and Contributions

The problem of AA is a significant topic in higher education, and it will receive more attention in the near future. The work presented in this paper aims at improving the practice of AA by proposing a model of an automated AA system, as well as one of the key components of such a model namely, a method for detecting student groups (clusters) based on the shared study paths. Study path is considered to be a sequence of courses a student has taken and completed during his/her studies. The paper presents the need for and the importance of proper AA. The paper makes two (connected) proposals: one for clustering students based on their study paths (to group students according to their study profiles in order to better advise them about future course selection), and the other for a model of automated AA system. The underlying approach is based on a clustering method (robust prototype-based clustering) which identifies groups of studentsâ study profiles. Moreover, the paper proposes a system architecture toward the development of an automated system for the academic advising process. There is no prototype of such a system. However, the discussion and comparison of the steps of the manual and the proposed automated academic advising process are presented.

The contribution of this work consists of a literature review on academic advising, the proposed clustering approach for student group profiling and the design of an architecture for an automated academic advising system.

Author’s contributions

The author of this thesis is the main author of this paper. She presented a literature review on academic advising, proposed a model of automated academic advising using clustered analysis, and described an architecture and model of the system, The data and the method of proposed system with the initialization for sparse data were performed by the third author. The second author described the data and the method of the proposed system, as well as described advantages of using clustering. Moreover, the main author presented this paper at 9th International Conference on Computer Supported Education conference (CSEDU), Porto, Portugal, 2017.

6.2.3 Article PVI: Smart Educational Process Based on Personal Learning Capabilities

This article was published in the proceedings of the 9th annual International Conference on Education and New Learning Technologies Conference (EDULEARN17), pages 6316–6324.

Objectives

The idea of this paper is to explain why it is important to have smart educational processes based on personal learning capabilities and to review the literature on different learning systems that detect learning progress. The paper also describes a model of the smart system which could be used for basic education and for adult education, which will personalize learning gaps according to incorrect answers.

Findings and Contributions

This paper presents a literature review on e-learning to see how different systems detect users' learning progress. Based on the survey, a method which takes into account a person's knowledge background and rate of material compatibility is presented. Authors provided an equation with which it is possible to check if the learner mastered the topic well and which will help in the elimination of knowledge gaps. To identify mastery, a person must pass a number of tasks at a specific threshold. The system should count false answers and detect knowledge gaps, take them into account, checking the relationships between topics in the Ontology. System will analyze where learner is making mistakes and find closely related topics which he or she has to study more. The system uses internal analysis and personalized testing. Also, an architecture of proposed system is presented. This approach is follows the style "test by loop". Periodic testing is the only certain way to ensure that no large knowledge gaps emerge.

Author's contributions

The author of this thesis is the main author of this paper. She presented a smart educational process description, system model proposal, method, and architecture. The third author presented the literature review, revised the paper, and helped with the introduction and conclusions. The main author presented this paper virtually at the 9th annual International Conference on Education and New Learning Technologies Conference (EDULEARN17).

6.3 Development of adaptive learning game

This section presents a special case about adaptive vocabulary learning for late talkers.

6.3.1 Article PIII: Adaptive vocabulary learning environment for late talkers

This article was published in the proceedings of the 8th International Conference on Computer Supported Education conference (CSEDU), 2016, pages 321–330.

Objectives

The focus of this work is in the area of facilitation learning techniques for children with an early language delay, as well as in proposing how to make LE more adaptive, taking into account that late learners require a certain level. The main aim is to recognize the personal qualities of the learner in building his or her model and to optimize the learning process using the personal learnerâs model for studying content creation. The idea behind this publication is to provide children with an early language delay with an adaptive way to train their vocabulary, taking into account the personality of the learner.

Findings and Contributions

The paper describes an idea of developing an adaptive vocabulary learning system for late talkers who have difficulties in learning language due to individual reasons.The system will be developed as a mobile learning application. The specific idea is that the functionality of the system is based on different similarity factors (visual, phonetic, and semantic) of the learning objectives, which are words. The combination of weighted similarity measures computes the complexity of a single task. Adaptivity of the system results in decreasing or increasing complexity of the presented tasks while also experimenting with the various dimensions that make up the similarity of the concepts. The general underlying assumption is that distinguishing similar concepts is harder than choosing from images of less similar concepts. The valuable part of the paper is the discussion of the three similarity dimensions and how these can automatically be measured. According to the user's feedback, the weights of the factors and similarity distance are adjusted to modify the level of difficulty in further iterations. The model of the system is designed to attempt to retrieve knowledge about the learners by recognition of aspects that are difficult for them. This approach may have serious consequences affecting the child's future life in a wide sense.

Author's contributions

The author of this thesis is the main and corresponding author of this publication. She searched and conducted the literature review on research related to AL and developed a model of AL system for children with difficulties, taking into account their personal qualities of concepts and perception. Also, the author described all concept similarities and part of the system architecture and adaptation logic. The second author also worked on the adaptation logic description and future work description. The third author of this paper described late talkers and took part in the proofreading and description of the concept similarities. Moreover, the main author presented this paper at 8th International Conference on Computer Supported Education conference (CSEDU), Rome, Italy, 2016.

6.3.2 Article PVII: On developing adaptive vocabulary learning game for children with an early language delay

This article will be submitted to the GSTF Journal on Computing (JoC), 2017.

Objectives

Late talkers perceive word categories differently than their typically developing peers. The task of this research is to explore how this could be investigated through game-based LE, while also creating a motivational way for increasing vocabulary size for children with an early language delay. The game-based tool should serve as research experiment and vocabulary learning.

Findings and Contributions

This article presents the need of adaptive game-based LE for children who are slow learners. The architecture of the system was presented, as well as all mechanisms of three main factors which could influence children vocabulary learning, are described. The ontology of animals was presented for semantic similarity counting and ranking. The artifact is developed. A shallow narrative component was created to serve the purpose and structure of research experiments that contain three phases of recognizing novel items, learning them, and testing the gained knowledge. This experiment is based on taking into account the semantic similarity of the words.

Author's contributions

The author of this thesis is the main author of this paper. She presented a review of AL systems for children with disabilities and the main descriptions of the system process, which is concentrated on semantic similarity words. She proposed a mechanism of filtering, learning, and testing. She wrote the main part of the article. The second author was responsible for the development of the game-based environment and also worked on the literature review. The third author also worked on the description of mechanisms and architecture of the whole system. The fourth author revised paper and structured it. The main author will submit this article to the journal GSTF Journal on Computing (JoC).

6.4 Summary of contributions

The results of the papers have contributed to three main aspects related to AL (i.e., theoretical development of ILSS, academic advising, knowledge gaps detection, and developing an adaptive vocabulary LG). The adaptation mechanism and adaptation technique used in the papers are presented in Table 1. The main results are summarized in Table 2.

In Table 2, the positioning of obtained results to each research questions is presented. Decision support system for LMS and automated AA present the answer to Research Question 3 of how to recommend a learning path for the learner, based on individual particular skills and preferences. The decision support system recommends a list of courses to obtain the desired specialization, based on the learnerâs background. The AA process uses robust clustering to find similar learning paths and recommend them to learners. Adaptive systems as enablers of feedback present the answer for Research Question 1 of how to make LEs adaptive and what kind of methods, techniques, and approaches could be used for that; and presents the answer for Research Question 2 of what kind of feedback to use while creating AL systems with high levels of adaptivity. For a higher level of adaptivity, the system should have feedback in the form of recommendations, progress tracking, and reports, with updated information according to usersâ actions and diagnoses of users' mistakes. ILSS also presents the answer to Research Question 1. Adaptive vocabulary LE presents the answer to Research Question 5 with the creation of a tool for adaptive recommendation of appropriate learning material for slow learners, according to their difficulties in the learning process. In this case, the similarity of words has been taken into account. Smart educational processes, based on the personal learning capabilities model, answers Research Question 4 of how to detect learning gaps and take them into account. Learning gaps could be taken into account by analysis of learnersâ mistakes and finding connected topics which may cause the difficulties.

The most essential findings related to the research are as follows:

- The model, which shows the relationship between adaptive system features and feedback types.
- The model of RS, which proposes further professional education based on the learner's current skills.
- The model of anAL system for slow learners, which takes into account multicriterial factors. The filtration method based on Semantic similarities was further developed.
- The model of an innovative learning ecosystem based on Semantic annotation with supporting modules.
- The model of automated academic advising process based on courses that were taken by other similar students.
- The model which allows to detect insufficient level of knowledge of particular topic.

TABLE 1 Summary of the results of included articles

LE or LMS	Adaptation mechanism	Adaptation technique
[PI] LMS	Providing recommendations based on known user data, personal preferences, and prediction of future specialization according to learners' knowledge background	SW technology
[PII] Game-based LEs	Review of technologies, features, and levels of adaptation	Feedback types, used in AL systems
[PIII], [PVII] Game-based LE	Concept similarity detection, which could influence on learning speed; personalization of complexity tasks	Review of different methods for visual, phonetic, semantic similarities, and measuring; manipulating coefficients for PL
[PIV] Ecosystem of ILSS	Modules, which take into account learners' skills, preferences, labor market changes	SW technology (Semantic annotation)
[PV] University (higher education)	Recommendation of learning path, based on evidence from the actual study paths of other similar students	LA (robust clustering)
[PV] Learning tool	Knowledge gap detection by analyzing learners' mistakes and finding relations between learning tasks (topics)	Knowledge Engineering (SW technology) and Watson Technologies

TABLE 2 Positioning of obtained results to each research questions

Obtained Results	Description	Article no:	Research question
a) Decision support for LMS	Models of course recommendations, based on knowledge background and preferences	[PI]	RQ3
b) Adaptive systems as enablers of feedback	Overview of recent English language learning game-based environments, discussed types of feedback in different adaptive system levels	[PII]	RQ2, RQ1
c) Adaptive vocabulary LE	Model of AL environment for late talkers, which consider word similarity and changes the level of complexity according to learners' skills	[PIII], [PVII]	RQ5
d) Intelligent Learning Support System	Ecosystem of ILSS, which takes into account learners' skills	[PIV]	RQ1
e) Automated academic advising	Supporting students with guidance (study path recommendation) utilizing prototype-based clustering	[PV]	RQ3
f) Smart educational process based on personal learning capabilities	Model of the system, which takes knowledge gaps into account	[PVI]	RQ4

7 CONCLUSION

This dissertation work presents research in the field of personalized and adaptive learning, utilizing technologies of intellectual data analysis and standards of the W3C. This is a solution of the current research topic, according to the main research goal. Research in the field of PL is in great request because of the growing demand for this technology. AL is an approach that takes into account the individual abilities and needs of the learner. With the active development of information technologies, LEs are finding increasing application in the field of education, which allow implementing the ideas of AL in practice. The use of adaptive technologies presupposes the integration of information and pedagogical technologies that provide interactivity between the subjects of education and the productivity of the student's learning activity with the use of new information technologies that provide adaptability in the educational process. The studied material is provided to the student at some point in time, taking into account his or her accumulated knowledge, academic performance, and experience. The adaptation process is realized in learning, while an adaptive presentation of course materials takes place.

It can be stated that AL is not just another approach to the formation of the learning process, but it also provides higher learning outcomes that can be measured quantitatively, as well as a personal trajectory of learning and student involvement in the learning process.

Based on the conducted studies of the thesis, one can draw the following conclusions:

– The challenges of LEs are revealed. One of the main problems is the lack of individual approaches to teaching students. The methods and techniques of developing the tools for the personalization of LEs are explored.
– According to the analysis of adaptive levels and feedback, which related to these levels, it is possible for developers to construct the specific desired adaptation level for learning system, including specific needed feedback for this level.
– A lifelong learning ecosystem has been proposed that will take into account

the knowledge and skills of learners so that a person can be in demand on the labor market, and also to fill their knowledge gaps. This model is based on semantic annotation. This ecosystem could be used for the creation of a lifelong learning process with future career development.

- An algorithm for recommending courses to obtain the desired specialization, based on the SW technologies, is proposed.
- A method for automatically recommending a learning path for students based on the similarity of teaching students as academic advising based on the robust clustering method is proposed. This method helps to determine the groups of actual profiles of students given their studies.
- A method to adapt an LG for young children who are learning words is proposed. This method takes into account the similarity of words in the factors of semantics, visualization, and pronunciation. When changing these indicators, the LG can adapt the process of learning words to each individual child.
- A training game for children has been developed that takes into account the semantic distance of words to check whether semantics affects the perception of words during instruction for children.

8 LIMITATIONS AND FURTHER RESEARCH

The studies carried out within the framework of this dissertation have certain limitations. This appears to be due to the fact that the ontological approach of KBs creation is not used in many subject areas. For example, the ontology described in article **PI** has a description of the students, disciplines and specializations, but specializations are described only by sets of disciplines (courses), and not by a set of competences (knowledge and skills) that student should possess after successfully completing of the program. This problem is described in detail in article **PIV**. It should be noted that the ontology implemented in articles **PIII** and **PVII** have limitations in the subject area that describes animals which are classified according to the biological principle. This clearly imposes certain limitations on the use of this KB in the learning process, because it describes only one subject area.

In the future, an extension of subject areas descriptions with an ontological approach is expected, which will greatly expand the possibilities for teaching children. As an example, the KB will contain not only the description of animals, but also other subject areas necessary for enriching the vocabulary of children. Accordingly, the use of a multidisciplinary ontology will broaden the competencies described in article **PIV**. Also, the question of keeping ontologies as recent one should be studied properly, so the knowledge which ontologies present will be up to date.

In paper **PII**, the relationship between the feedback types and areas of English learning is shown. The limitation of this work is that only 23 LGs designed to teach English were analyzed in this study. Future work has been planned to incorporate other AL games in different disciplines to analyze adaptation levels and feedback types used in those areas.

It is also limitation connected with the fact that there is only models, not full systems for automating the process of recommendations and determining the academic level of knowledge and skills have been developed and used in this work. Paper **PIV** explains why it will be better to use annotations based on knowledge and skills for description of learners, courses, study programs, jobs, etc. But it is not describing the development process of the whole system and the is no explanations of how to keep ontologies, which will be used in the system, as an up to date knowledge. It means that adaptation is needed not only for the learning process, but for the system as well.

Paper **PIV** explains why using annotations based on knowledge and skills for a description of learners, courses, study programs, jobs, among other outcomes, will improve understanding of this topic. However, Paper **PIV** does not describe the entire development process, and it does not explain how to maintain ontologies within the system. Adaptive measures must be used in both the learning process and the system itself to ensure accurate results are obtained. It is no explanations of how to keep ontologies, which will be used in the system, as up-to-date knowledge. It means that adaptation is needed not only for the learning process but for the system as well.

In paper **PI**, **PV** and **PVI**, a model of recommending study path according to learners' preferences, capabilities, and learning process is provided. The description of architecture is made, and a description of the recommendation (advising) process is also given while addressing future development of the recommendation and advising mechanism. It is intended to develop software tools that will use approaches and models for adapting the learning process described in the articles. To complete the full development cycle, the created applications will be tested, and necessary corrections will be made after the testing process is completed.

The development of adaptive LG in papers **PIII** and **PVII** was limited to the game, which only takes into account the semantic similarities between words for future testing on children, to detect whether semantics influences their learning speed and ability to recognize specific words among other similar words. This testing should be used first with normal children to detect how they learn, and then with slow learners for future comparison of children's behavior. After that, according to the results, other similarity factors could be added, changing their weights according to the child's results, to personalize the system for each child.

YHTEENVETO (FINNISH SUMMARY)

Oppimisympäristöjen personoitu adaptaatio

Nykyisten oppimisympäristöjen ongelmana on se, että ne lähestyvät yleensä kaikkia oppilaita samalla tavalla tarjoamalla oppilaille samanlaisia tehtäviä, materiaaleja ja suoritusaikaa. Oppilaat ovat kuitenkin yksilöitä, joiden taidot, kyky omaksua tehtäviä ja tavoitteet vaihtelevat. Suosittelujärjestelmät ovat yleistyneet oppimisjärjestelmissä, mikä on yksi osa personoitua oppimista. Tämän on mahdollistanut matemaattisten menetelmien, kuten koneoppimisen ja tiedonlouhinnan, hyödyntäminen ja teknologisten standardien käyttö (W3C). Matemaattisten menetelmien ja uuden ohjelmistoteknologian ansiosta voidaan toteuttaa järjestelmä, joka määrittelee opiskelijan nykyisen taitotason, seuraa oppimisprosessia ja suosittelee parempia oppimismateriaaleja. Palaute-, ohjaamis- ja suosittelujärjestelmien analysointi tukee tätä ajatusta. Tutkimukseni havainnollistaa kuinka oppimisympäristöjä voidaan kehittää mukautumaan erilaisiin käyttäjiin, perustuen heidän tietoihinsa, käyttäytymiseensä, ominaisuuksiinsa ja tavoitteisiinsa. Tässä työssä esitellään ensin oppimisympäristön malli, joka perustuu opiskelijan tietojen ja taitojen annotointiin. Tämä malli havainnollistaa opiskelijan elinikäistä oppimisprosessia. Sen jälkeen työssä keskitytään personoitujen arviointi-, suosittelu- ja ohjaamisjärjestelmien kehittämiseen. Lopuksi, työssä kehitetään mukautuva oppimispeli, joka huomioi oppilaiden erilaisen kyvyn havaita sanoja.

REFERENCES

Ahmed, F., Rashid, S. & Khan, R. 2013. A proposed model of e-learning management system using semantic web technology.

Attwell, G. 2007. Personal learning environments-the future of elearning? Elearning papers 2 (1), 1–8.

Bailey, G. D. 1992. Wanted: A road map for understanding integrated learning systems. Educational Technology 32 (9), 3–5.

Bellotti, F., Ott, M., Arnab, S., Berta, R., de Freitas, S., Kiili, K. & De Gloria, A. 2011. Designing serious games for education: from pedagogical principles to game mechanisms. In Proceedings of the 5th European Conference on Games Based Learning. University of Athens, Greece, 26–34.

Berners-Lee, T., Hendler, J., Lassila, O. et al. 2001. The semantic web. Scientific american 284 (5), 28–37.

Berners-Lee, T. 2005. Putting the web back in semantic web. ISWC, Galway.

Bittencourt, I. I., Isotani, S., Costa, E. & Mizoguchi, R. 2008. Research directions on semantic web and education. Interdisciplinary Studies in Computer Science 19, 59–66.

Bouquet, P. & Molinari, A. 2012. Semantic technologies and e-learning: Towards an entity-centric approach for learning management systems. Journal of e-Learning and Knowledge Society 8 (2).

Breuer, J. S. & Bente, G. 2010. Why so serious? on the relation of serious games and learning. Eludamos. Journal for computer game culture 4 (1), 7–24.

Capuano, N., Gaeta, M., Orciuoli, F. & Ritrovato, P. 2009. On-demand construction of personalized learning experiences using semantic web and web 2.0 techniques. In Advanced Learning Technologies, 2009. ICALT 2009. Ninth IEEE International Conference on. IEEE, 484–488.

Carapina, M., Staničić, O. & Nožica, B. 2013. The semantic web technologies and their integration within e-learning systems: an overview. In ICERI 2013.

Castellanos-Nieves, D., Fernández-Breis, J. T., Valencia-García, R., Martínez-Béjar, R. & Iniesta-Moreno, M. 2011. Semantic web technologies for supporting learning assessment. Information Sciences 181 (9), 1517–1537.

Castro, F., Vellido, A., Nebot, À. & Mugica, F. 2007. Applying data mining techniques to e-learning problems. Evolution of teaching and learning paradigms in intelligent environment, 183–221.

Chakraborty, S. & Mittal, R. 2010. Semantic web technology: A strategic approach to intelligence. International Journal of Computer Applications 5 (9), 33–37.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U. & Thüs, H. 2012. A reference model for learning analytics. International Journal of Technology Enhanced Learning 4 (5-6), 318–331.

Chen, C.-M. 2008. Intelligent web-based learning system with personalized learning path guidance. Computers & Education 51 (2), 787–814.

Comunale, M. A. 2017. Getting Into the Game: An Explanatory Case Study to Examine the Experiences of Faculty Incorporating Digital Game Based Learning in Higher Education. Drexel University. Ph. D. Thesis.

Cristea, A. 2004. Adaptive and adaptable educational hypermedia: Where are we now and where are we going. Proceedings of Web-based Education, Feb, 16–18.

Devedžic, V. 2006. Semantic web and education, Vol. 12. Springer Science & Business Media.

Dias, S. B. & Diniz, J. A. 2014. Towards an enhanced learning management system for blended learning in higher education incorporating distinct learners' profiles. Journal of Educational Technology & Society 17 (1).

Ellis, R. K. 2009. Learning managament systems. Alexandria, VI: American Society for Training & Development (ASTD), 1–7.

Ferguson, R. 2012a. Learning analytics: drivers, developments and challenges. International Journal of Technology Enhanced Learning 4 (5-6), 304–317.

Ferguson, R. 2012b. The state of learning analytics in 2012: A review and future challenges. Knowledge Media Institute, Technical Report KMI-2012-01.

García-Barrios, V. M., Mödritscher, F. & Gütl, C. 2005. Personalisation versus adaptation? a user-centred model approach and its application. In Proceedings of the International Conference on Knowledge Management (I-KNOW), 120–127.

Ghazal, S. & Singh, S. 2016. Game-based language learning: Activities for esl classes with limited access to technology. ELT Voices- International Journal for Teachers of English 6, 1–8.

Gilhooly, K. 2001. Making e-learning effective. Computerworld 35 (29), 52–53.

Gray, G. 2014. A review of psychometric data analysis and applications in modelling of academic achievement in tertiary education. Journal of Learning Analytics 1 (1), 75–106.

Gruber, T. R. 1995. Toward principles for the design of ontologies used for knowledge sharing? International journal of human-computer studies 43 (5-6), 907–928.

Guangzuo, C., Fei, C., Hu, C. & Shufang, L. 2004. Ontoedu: a case study of ontology-based education grid system for e-learning. In GCCCE2004 International conference, Hong Kong, 1–9.

Halimi, K., Seridi-Bouchelaghem, H. & Faron-Zucker, C. 2014. An enhanced personal learning environment using social semantic web technologies. Interactive Learning Environments 22 (2), 165–187.

Harper, B., Agostinho, S., Bennett, S., Lukasiak, J. & Lockyer, L. 2005. Constructing high quality learning environments using learning designs and learning objects. In Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on. IEEE, 266–270.

Heiner, C., Heffernan, N. & Barnes, T. 2007. Educational data mining. In Supplementary Proceedings of the 12th International Conference of Artificial Intelligence in Education.

Jere-Folotiya, J. 2014. Influence of grade one zambian teachers and graphogame on initial literacy acquisition: Lusaka district. Jyväskylä studies in education, psychology and social research; 0075-4625; 503.

Kankaanranta, M., Koivula, M., Laakso, M.-L. & Mustola, M. 2017. Digital games in early childhood: Broadening definitions of learning, literacy, and play. In Serious Games and Edutainment Applications. Springer, 349–367.

Kankaanranta, M. 2007. Digital games and new literacies. Finnish Reading Literacy. When Equality and Equity Meet, 281–307.

Kansanaho, J. 2014. Usuko-new generation school. In EdMedia: World Conference on Educational Media and Technology. Association for the Advancement of Computing in Education (AACE), 802–806.

Kats, Y. 2010. Learning Management System Technologies and Software Solutions for Online Teaching: Tools and Applications: Tools and Applications. IGI Global.

Khozooyi, N., Malekhoseini, R. et al. 2012. Ontology-based e-learning. IRACST-International Journal of Computer Science and Information Technology & Security (IJCSITS) 2 (4).

Kiili, K. 2007. Foundation for problem-based gaming. British journal of educational technology 38 (3), 394–404.

Kiili, K. 2010. Call for learning-game design patterns. Educational Games: Design, Learning, and Applications.

Kirriemuir, J. & McFarlane, A. 2004. Literature review in games and learning.

Kobsa, A. 2001. Tailoring privacy to usersâ needs. User Modeling 2001, 301–313.

Kohlhase, A. & Kohlhase, M. 2008. Semantic knowledge management for education. Proceedings of the IEEE 96 (6), 970–989.

Lee, M.-C., Tsai, K. H. & Wang, T. I. 2008. A practical ontology query expansion algorithm for semantic-aware learning objects retrieval. Computers & Education 50 (4), 1240–1257.

Lytras, M. D., Pouloudi, A. & Korfiatis, N. 2003. An ontological oriented approach on e-learning. integrating semantics for adaptive e-learning systems.

Markellou, P., Mousourouli, I., Spiros, S. & Tsakalidis, A. 2005. Using semantic web mining technologies for personalized e-learning experiences. Proceedings of the web-based education, 461–826.

Nat, M., Walker, S., Bacon, L. & Dastbaz, M. 2010. Designing personalisation in lams. In Proceedings of the 2010 European LAMS & Learning Design Conference, 120–127.

Ndongfack, M. N. 2016. Retracted: Design and development of a personal learning environment for corporate self-regulated learning. Journal of Computer and Communications 4 (04), 1.

Ouf, S., Ellatif, M. A., Salama, S. E. & Helmy, Y. 2017. A proposed paradigm for smart learning environment based on semantic web. Computers in Human Behavior 72, 796–818.

Oxman, S., Wong, W. & Innovations, D. 2014. White paper: Adaptive learning systems. Integrated Education Solutions.

Pahl, C. & Holohan, E. 2009. Applications of semantic web technology to support learning content development. Interdisciplinary Journal of E-Learning and Learning Objects 5.

Pardo, A. & Teasley, S. 2014. Learning analytics research, theory and practice: widening the discipline. Journal of Learning Analytics 1 (3), 4–6.

Prensky, M. 2003. Digital game-based learning. Computers in Entertainment (CIE) 1 (1), 21–21.

Reigeluth, C. M. & Garfinkle, R. J. 1994. Envisioning a new system of education. Systemic change in education, 59–70.

Reigeluth, C. M. 1997. Educational standards: To standardize or to customize learning? Phi Delta Kappan 79 (3), 202.

Richardson, U. & Lyytinen, H. 2014. The graphogame method: The theoretical and methodological background of the technology-enhanced learning environment for learning to read. Human Technology 10 (1).

Rogers, T. 2015. Critical realism and learning analytics research: epistemological implications of an ontological foundation. In Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. ACM, 223–230.

Romero, C. & Ventura, S. 2010. Educational data mining: a review of the state of the art. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 40 (6), 601–618.

Rumetshofer, H. & Wöß, W. 2003. Xml-based adaptation framework for psychological-driven e-learning systems. Educational Technology & Society 6 (4), 18–29.

Saarela, M. 2017. Automatic knowledge discovery from sparse and large-scale educational data: case finland. Jyväskylä studies in computing 262.

Santally, M. I. & Senteni, A. 2005. Adaptation models for personalisation in web-based learning environments. Malaysian Online Journal of Instructional Technology 2 (1), 1–23.

Shum, S. B. & Ferguson, R. 2012. Social learning analytics. Journal of educational technology & society 15 (3), 3.

Siemens, G. 2012. Learning analytics: envisioning a research discipline and a domain of practice. In Proceedings of the 2nd international conference on learning analytics and knowledge. ACM, 4–8.

Siemens, G. 2013. Learning analytics: The emergence of a discipline. American Behavioral Scientist 57 (10), 1380–1400.

Sonwalkar, N. 2005. Technology for adaptive learning: from one size-fits-all to adaptive individualization, educause center for applied research (ecar). Res. Bull, 1–11.

Stojanovic, L., Staab, S. & Studer, R. 2001. elearning based on the semantic web. In WebNet2001-World Conference on the WWW and Internet, 23–27.

Stracke, C. M. 2012. Learning innovations and learning quality: Relations, interdependences, and future. The Future of Learning Innovations and Learning Quality. How do they fit together, 13–25.

Sudarshan, M., Hegde, M. K. S. & Arun, M. L. 2012. Intelligent curriculum for learning analytics. International Journal of Science and Engineering Applications (IJSEA) 1, 4–7.

Terziyan, V., Golovianko, M. & Shevchenko, O. 2015. Semantic portal as a tool for structural reform of the ukrainian educational system. Information Technology for Development 21 (3), 381–402.

Venable, M. 2011. Adaptive learning technology: An introduction.

Wang, H.-C. & Hsu, C.-W. 2006. Teaching-material design center: An ontology-based system for customizing reusable e-materials. Computers & Education 46 (4), 458–470.

Watanabe, C., Kanno, G. & Tou, Y. 2012. Inside the learning dynamism inducing the resonance between innovation and high-demand consumption: A case of japan's high-functional mobile phones. Technological Forecasting and Social Change 79 (7), 1292–1311.

Watanabe, C., Naveed, K. & Neittaanmäki, P. 2017. Co-evolution between trust in teachers and higher education toward digitally-rich learning environments. Technology in Society 48, 70–96.

Weber, G. & Brusilovsky, P. 2001. Elm-art: An adaptive versatile system for web-based instruction. International Journal of Artificial Intelligence in Education (IJAIED) 12, 351–384.

Özyurt, Ö. & Özyurt, H. 2015. Learning style based individualized adaptive e-learning environments: Content analysis of the articles published from 2005 to 2014. Computers in Human Behavior 52, 349–358.

ORIGINAL PAPERS

PI

SEMANTICALLY ENHANCED DECISION SUPPORT FOR LEARNING MANAGEMENT SYSTEMS

by

Mariia Gavriushenko, Marja Kankaanranta, Pekka Neittaanmäki 2015

9th International Conference on Semantic Computing (IEEE ICSC 2015),
Anaheim, California, USA

Semantically Enhanced Decision Support for Learning Management Systems

Mariia Gavriushenko, Marja Kankaanranta, Pekka Neittaanmäki
Dept. of Mathematical Information Technology
University of Jyvaskyla, Jyvaskyla, Finland
gavrushik@gmail.com
marja.kankaanranta@jyu.fi
pekka.neittaanmaki@jyu.fi

Abstract—**This article focuses on proposed semantically enhanced model of decision support system for learning management system (LMS). The model is based on a survey of LMSs and various plugins used in these to improve educational process. Systems based on semantic technologies are capable of integrating heterogeneous data, flexibly changing data schemas, semantic search (using ontologies), and joint knowledge development. The knowledge base that was developed for the proposed system model is presented in an ontological form. Ontology-based applications limit the "fragility" of the software and increase the likelihood of its reuse. In addition, they profitably redirect the efforts previously focused on software development and maintenance of creation and modification of knowledge structures. In the proposed knowledge base, we developed the necessary rules for further recommendations of specialization and courses for users. These recommendations are based on users' data extracted from profiles and user preferences.**

Keywords-Learning Management System, Ontology, Recommendation system, Semantic Technology, SWRL-rules

I. INTRODUCTION

During recent years the use of Learning Management Systems (LMSs) has been increased [1]. LMS is a good example of learning with information technology adoption with the needed courses for studies. LMS meets the requirements of society, while taking into account monetary and time costs.

It is possible to develop a set of intelligent tools for LMS that take into account the use of hybrid platforms, semantic integration of existing web-based applications and content aggregation. These intelligent tools will include a variety of advanced technologies that make the content more relevant to learning environments, with respect to the user, and the learning process more innovative.

Intelligent tools are used for solving difficult semi-structured tasks such as grouping, focusing attention, and combinatorial search [2]. They are characterized by their ability to make autonomous decisions [3]. The main examples of intelligent tools are expert systems, fuzzy set systems, artificial neural networks, and genetic algorithms [4]. One of the expert system is the recommendation system with decision support. Recommendation systems (RSs) are programs that attempt to predict which items (from a range of movies, music, books, news, sites, courses, etc.) will be of interest to the user. They make use of some information from the users' profile.

In addition to e-commerce, RSs could be implemented also in the educational domain [5], [6], [7], [8]. In modern RSs, semantic technologies are often used. This trend is based on the fact that the semantics assumes a standard syntax of data, Resource Description Framework, which is a W3C specification for standardizing a metadata model and language. The framework uses resources as operating elements and properties as the interactions between them. Any resource can be specified using universal resource identifier (URI), and the predicate between the subject and object is standing as a property, forming together a statement about interaction. [9] There is a standard way to describe the properties of the data (RDF-schema), which has largely accelerated the algorithm of a RS. Researchers pay a lot of attention to the usage of semantic technologies in construction of intelligent systems because these technologies have great potential to reflect the relationships between objects and concepts, and the description is thus close to natural language.

People want to be specialists in a specific field. And the option to become a specialist in a field motivates the need to create the most suitable learning path for a future profession.

In this study, the focus is on developing an intelligent tool for the design of a Learning Management System. The tool will be based on a study of semantically enhanced decision support while building a model of RS for LMS users. RS will function as a study adviser indicating which courses are suited best for the users while taking into account the users' background and preferences in future specializations. The system can suggests the kinds of courses necessary to study in addition to courses that the users have already completed.

This paper is structured as follows. Section 2 opens a discussion about semantic web technologies and ontology-driven systems. Section 3 presents general information about LMSs and about the need to use RS within the system. Section 4 presents a literature review about semantic technologies which can be used for developing LMSs. Section 5 explains our work, providing specifications and requirements and suggested architecture for LMS with a decision support tool, as well as details of building ontology and rules for using in RS.

Section 6 concludes the paper by discussing future extensions and improvements.

II. Semantic-Web-Oriented Technologies

Semantic Web contains data which is defined and linked in such a way that it can be used by computers not only for display but also for automatic processing, integration and reuse by different applications, as well as for extracting information and getting new knowledge [10].

Semantic Web offers a new approach to managing information and processes. The fundamental principle of this approach is the creation and use of semantic metadata. Metadata can exist on two levels. On the one hand, metadata can describe a document, for example, a web-page or a part of a document such as an item. On the other hand, it can describe objects in documents, such as a person or company. The most important feature of semantic metadata is that it can describe the content of the document or objects in the document [11].

Using semantic approach, the system is able to understand where the words are equivalent, which can increase the efficiency of the reporting method. Instead of searching for a linear list of results, the results can be grouped by a meaningful criterion.

In addition, it is also possible to use semantics to integrate information from all relevant documents, removing redundancy and tabulation, if necessary. It is possible to draw conclusions from the knowledge available for creating new knowledge. Semantic Web allows improving the associated metadata with e-learning materials as well as extension of the existing options for e-learning states [12].

At the heart of all Semantic Web applications there are the ontologies, which play an important role for various applications in knowledge representation, processing, sharing and reuse [13]. Ontology-based semantics are utilized mostly for improving the analysis of information in unstructured documents. Domain ontology plays a central role as a resource to structure the learning content. Ontology was originally defined, in 1993, by Gruber: "Ontology – is explicit and formal specification of a conceptualization of a particular interest" [14]. Ontology (in computer science) is an attempt to formalize comprehensive and detailed knowledge of domain with the help of a conceptual schema. Typically, such schema consists of a data structure containing the relevant classes of objects and their relationships and rules (theorems, limits) in the domain. Ontologies play a prominent role in the processing of and sharing in the knowledge-based web. Information exchange can take place amongst people and applications.

Usability of ontology is enhanced by software tools that provide a general domain-independent support for ontological analysis. A number of ontological analysis tools support editing, visualization, documentation, import and export of ontologies in different formats, their presentation, association and comparison. Protégé and its commercial analogue TopBraid Composer are among the existing solutions today for working with Semantic Web.

In modern LMSs, ontologies are useful because they specify conceptualization of a specific domain in terms of concepts, attributes and relationships.

III. Learning Management Systems

Learning Management System allows improving one's level of education and provides a range of training programs. LMS helps in developing new opportunities in the learning domain. Students can get quick feedback, and teaching methods can be individual for each learner.

According to Ellis and Calvo [15], LMSs are programs which are used for the storage, management and delivery of education courses or training programs. A sturdy LMS might perform automation and centralization of administration, use self-guidance and self-services, gather and transfer learning content quickly, fund training initiatives on a scalable web-based platform, support standards and portability, personalize content, and allow knowledge reuse [16]. Most of LMSs have commercial software licenses, but there are also open-source LMSs [17].

LMSs have the main following functions [18]:

- Content delivery and development;
- Class and user management;
- Formative and summative assessment;
- Asynchronous and synchronous communication.

The organization and management of LMSs, its teaching activities and content can be simplified with collaborative intelligent-technology software modules. Development of such modules is aimed to improve students' performance. Data from the users' profiles is employed and users' actions are taken into account.

There are many theories and practices proposed to improve Web-based learning environments [19]. For example, content management systems helping students and teachers to implement instructional strategies; digital repositories that help sharing and reusing educational resources; ontologies that help to build students' profiles for achieving a personalized learning level.

Today, there is a vast range of LMS courses, providing the user with a variety to choose from. However, in order to find the courses that will meet the users' needs concerning their desired profession, users need to spend a lot of time going through the entire courses' descriptions. This will also help them find out about the learning outcomes of each interesting course. It is better to use a decision support system and filter the rates in different conditions. This will save time for the potential subscriber of LMS.

IV. Semantic Technologies For LMSs

Semantic Web technology has brought a set of foundational technology components, both at the conceptual level and at the level of prototypes and software, among other assets, which include ontology engineering methodologies, standardized ontology languages, ontology-engineering tools, and other infrastructure like APIs, repositories and scalable reasoners. The availability of this technology makes the development of adaptive systems or services in the paradigm of Semantic web possible.

Earlier studies have proposed diverse systems for developing LMSs based on semantic web technologies. The most prominent of these systems are intelligent tutoring systems [20], [21], e-LMSs [12], [22], [23], [24], multi agent systems which capture users' behaviors [25], personalized RSs [26], [27], adaptive LMS [28], [29], and Ontology-based LMS [30], [31], [32].

Artificial intelligence techniques with ontologies and SWRL, for example, can be used by intelligent tutoring systems to build a pedagogical model, which is accountable for selection and organization of teaching strategies and tactics for future communication with the students and for offering a personalized teaching [20]. These systems can be used to create practical course material in succession as a generalized model for e-learners [33] and to make recommendations for students using LMS. Also, such systems are reasonable for the development of e-learning courses and for the use of them by students. These systems could be based on XML-technologies (XML-schema and XSLT) with Jess [21].

In the model of LMS, which uses semantic technologies, learning activities and learning styles as teaching methods, course syllabi are often included, as they are more flexible for various training institutions. In this model, there are semantic relations between concepts, and useful information for searching and sequencing learning resources in LMS is provided [12]. Learning processes, search of learning materials, learner personality and learning approaches could be integrated into a semantic e-learning framework with the help of ontologies, semantic descriptions such as RDF/XML and XQuery (XML Query Language) [24]. Here we are dealing with new knowledge visualization and learning process modeling for building intelligent personalized e-learning [22].

A multi-agent system captures the users' behaviors and stores them for further analysis and reasoning. The system allows students to discuss any topics and assignments in an interactive online environment that simulates communication and transforms the collective knowledge in the generated Semantic Web, which is accessible not only by students who now take some courses but also by students who will take part in these courses in future. The system analyzes the pages that students have visited and keywords entered during the search process [25].

For building recommendations, earlier studies have provided requirements for formal e-learning scenarios. There is a need to extend the adaptive functions of LMS with the help of semantic web technologies for building RS. Santos and Boticaria [27] provided the three main requirements for developing semantic educational RS: recommendation model, open standards-based service-oriented architecture and graphical user interface for delivering recommendations.

Theoretical framework of adaptive LMS using semantic web technologies focuses on extracting knowledge from users' interactions, actions and behavior to transform and display it in an ontological representation, on finding the learner style, and on providing the work-flow according to learners' styles [28]. The system is to be used for recommendation of personalized work-flow for the learner [28]. Also, adaptive LMS may concentrate on adaptive management and reuse of e-learning resources. Approach that uses an evolution algorithm is presented in [29].

A main function of ontology-based LMSs is the creation of knowledge bases. Ontology technology can assist learners, teachers and developers in organizing, personalizing and reusing content [30]. Ontologies could benefit from improved content creation, configuration, management support, and reusability of experiment outcomes, enabling intelligent and personalized student support [30], [31]. Ontologies could also be used for describing courses and test questions based upon Mixed Diagnostic Tests, which use fuzzy logic and allow evaluation of the quality of mastery of the learning material [32].

Based on our analysis, we can say that there are theoretical models for making semantically enhanced LMS. There are also many systems which provide teaching strategies and learning materials and also systems for creating collective knowledge in LMS. Implementation of such adaptive systems into LMSs is extremely useful, because these systems can monitor and analyze user behavior. It is also important not only to obtain information about the course, person, materials, etc., but also to recommend courses for further study, taking into account the preferences of each involved user and knowledge about the users' background.

V. Intelligent LMS Decision Issues

A. Specifications for the designed LMS with a decision support tool

We propose a theoretical model of LMS with decision support. Taking into account the main requirements for building LMSs, there are some other extra proposed requirements for LMS with decision support tool. The system should also:

- Provide storage and facilitate manipulation of educational content, registration, authorization, and administration;
- Allow extraction of users' data for future analysis and create students' profiles according to extracted data;
- Support editing and saving user data;
- Support creation and editing of teachers' profiles;
- Provide users with a possibility to choose the desired specialization;
- Provide intelligent analysis of users' data (background, desired specialization, some user restrictions on desired courses);
- Store rules for future recommendations and for users' decision support;
- Store the rules in a flexible manner;

- Make recommendations to the users about their personal potential specialization, according to their background;
- Provide course recommendations for users, according to their background, desired specialization and restrictions on the courses.

Also, from the learner perspective, there is the need to include reusable requirements:

1. The storage and information management are needed for keeping the history of interactions between interested key users. This enables us to use RDBMS (relational database management system)-oriented systems during architecture creation. The RDBMS-provider can be vendor-independent; however, the relational nature is needed because of the former constraints from the domain field with the inter-linked data from learner, the history and learners' specializations.

2. The extraction of data and its analysis take us to the point of module aggregation in a separate service with links to LMS that are based on the principles of interoperability, modularity and interchangeability.

3. The storing of the created rules and potential specific extracted recommendations leads us to the point where the fore-mentioned domain-level needs to be represented as a knowledge base and a reasoner, in a form of separate modules in the architecture.

4. User interaction with LMS and the definition of the previous learner history (for example, the desired specialization) are followed up by a user interface module which is usually implemented by web application technologies such as ASPX or PHP.

B. Architecture of LMS services with decision support

Our web-service LMS is designed to be deployed on a private Cloud platform. This platform was chosen because it satisfies the requirements for availability, reliability and flexibility. Such web-service works for users on a 24-hour daily schedule. Virtual infrastructure will be deployed in the Cloud platform, which should have the necessary sub-components for the operations of LMS.

The regular life-cycle of the LMS requires a web server such as the Apache HTTP-server. Executable LMS code-base may be interpreted using PHP technology, which is connectable to Apache as a pluggable module.

Visitors' access the LMS Web User Interface by using HTTP-requests from supported browsers. Requests are transferred through the HTTP-protocol to the web server. The traditional model of the HTTP-protocol suggests that the interaction between clients and the server is based on the principle of independent transactions. The client sends a request to the server (or calls a specific page); the server in turn responds with a properly formatted response to the client (or with details of that page). HTTP-transactions are considered independent. This model does not force the server to store users' data in-between transactions; thus the performance does not change. A further HTML-page will be generated and sent to the visitor.

An RDBMS-database can be used for storing LMS-data. For example, an open-source, highly documented Oracle MySQL vendor-based solution might be chosen. The system input data will be user information consisting, in particular, of personal information that does not change over time: i.e. name, sex, date of birth and place of birth. This functions as a background, which includes the knowledge the user has received during education and the user preferences for specialization. Also MySQL can save registration data, data about courses, learning outcomes for them, and information about teachers, to mention a few.

The Apache Tomcat web server and Java Virtual Machine will be used for the deployment of intelligent service. Apache Tomcat allows running web applications; it contains a number of programs to self-configure.

The data will be extracted from a MySQL database and sent back to the intelligent services, which will put it in a knowledge base that is represented as ontology. The obtained data will be converted from JSON/XML (as a transfer format) to RDFS and stored in the ontology for SWRL-rules processing.

The reasoner has functions of an inference machine. It allows making direct queries to the data domain ontologies written in SPARQL.

The architecture of LMS with a decision support tool is presented below in Fig. 1.

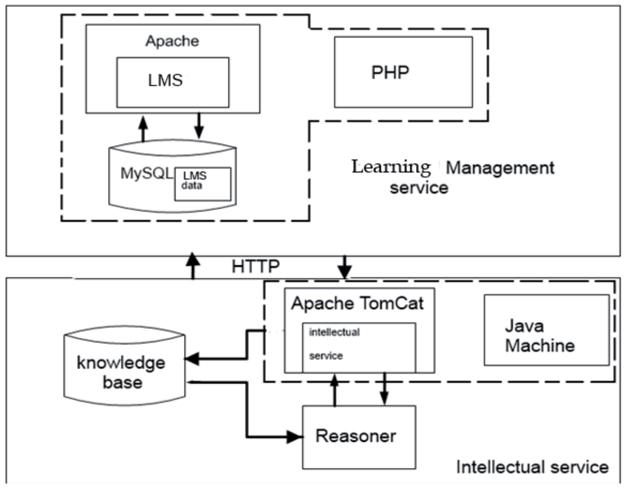


Figure 1. Architecture of LMS with a decision support tool

The designed web application is multi-component and includes:

- Administrative web interface, which acts as the management of LMS;
- Web-based user interface that provides user access to the LMS;
- MySQL database: the developed database stores low-level information: for example, information about courses, teachers, visitors, user authorization keys, users' background, learning outcomes, etc.;

- Knowledge base, represented ontologically. Access to the knowledge base will be by intelligent applications and machine of inference;
- Pellet reasoner. The knowledge base contains knowledge about the data domain, which includes information about the users of LMS, their background, their purpose, courses, teachers, etc.

The knowledge base will provide the ability to store data about users and process them with SWRL-established rules, which provide the means to display new facts of existing knowledge for further course recommendations for users. SWRL-rules include the data about users' background and their preferences and courses. With the help of these specific rules, the system can take into account users' background and preferences and offer new courses for training and future specialization.

The information about the recommendations on specific courses for the user will be extracted from the knowledge base with the help of SPARQL-queries, and it will be associated with the user's account that is persisted in the database. Then the user enters recommendations about the courses which might take into account his or her background and preferences.

C. *Universal system of authentification*

OAuth is a universal system of authentication. It is very useful in LMSs, because users do not need to spend time on registration and filling in some basic information about them. The schema using extraction data from social network profiles for the LMS user authentication/authorization is presented on Fig. 2.

The user visits the LMS website, fills in the login form and confirms the completion of the required fields. The application, which provides authentication/authorization of users in the LMS via the OAuth protocol, requests a token from Facebook. After a successful authentication, in response to an OAuth protocol request, the LMS application sends the requested token and a secret token. Then the application redirects the user to a page to confirm an agreement for providing access to the personal data.

Once the user has agreed to the use of personal data from the account in the social network account, Facebook sends a token and a secret token to the LMS. The tokens will be saved in the database and used in the future for each user request. The tokens are unique for each user. The user data obtained is converted from the JSON / XML format to RDFS and stored in the ontology for further use.

D. *Ontology design*

Protégé Ontology Editor was used to develop the ontology for LMS [34]. In this study, the ontology is provided for storing information about users, teachers, courses, specializations, and universities, among other things. Also, an ontology is needed for SWRL-rules formulation to further classify users into groups by their background. The ontology can be found at: http://nomad.dyndns.ws/magavriu/EMS.owl. The diagram of the ontology developed is presented below in Fig. 3.

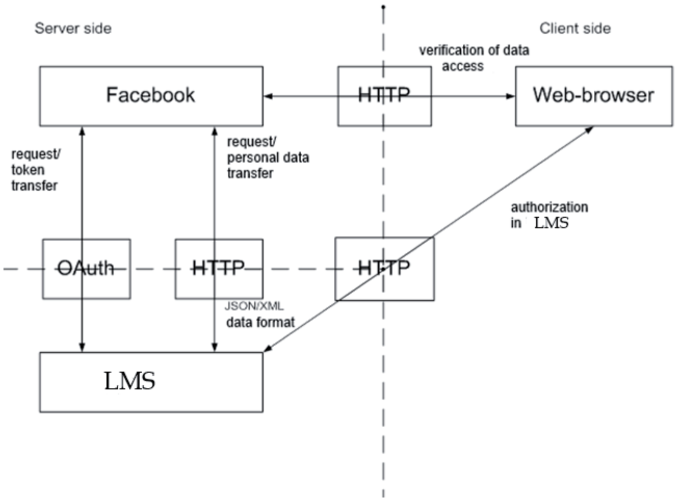


Figure 2. The schema of the authentication/authorization for LMS users using data extraction from their social network profile

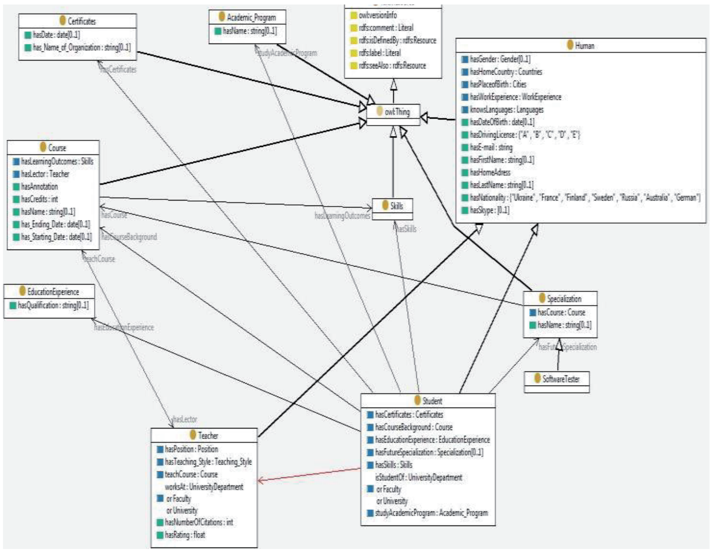


Figure 3. Diagram of the ontology for LMS

There is a need to distribute users into groups, depending on what data about the visitors of LMS is known. For classification of people into groups and for building links between these groups and their course groups, it is necessary to generate new knowledge in the ontology developed.

For the formulation of new knowledge in ontologies, we used SWRL (Semantic Web Rule Language) and Jena technologies. SWRL is a technology that helps to describe the mechanism for manipulating abstract objects in the domain area. The main advantage of SWRL is that it allows deriving new facts from existing approvals. SWRL-rules can be described using languages such as CLIPS or Prolog, which increases the efficiency of the description and processing [35].

SWRL was created by combining OWL DL, OWL Lite (axioms, facts) and Rule ML (rules), i.e. SWRL combines descriptive logic and production of external logic.

This technology is necessary for the following reasons:

- SWRL rules do not contain specific objects; they only refer to them. This makes it possible to apply the same rule to a number of groups of objects;
- SWRL Terms can be added to the OWL-description, i.e. included in the ontology;
- The process of writing and reading the rules is easier if we use a special language that is close to a human-language group.

SWRL-rules are needed to define possible future recommendations for the students. In line with this point, the concept of specialization will contain a set of courses for which recommendation can be used. The system will produce such recommendation based on the user input, i.e. "to be a software manager you need to take these courses". The ontology in question contains specialization details and holds knowledge about the domain (set of courses for the specialization).

Then the user describes or specifies which set of courses he or she has already obtained or whether the recommendation did not suit the user needs. Finally, using the newly-specified user input, the system recommends updated courses for further study. The recommendation is an entity generated by the Learning Management System - Intellectual Service.

An example of a SWRL-rule in the developed ontology is: "if student has course background "Database and Knowledge Management", "Software testing and QA engineering"", then the future specialization of this student might be "Software Tester". The code is presented below.

```
Student(?s), hasCourseBackground
(?s, DatabaseAndKnowledgeManagement),
hasCourseBackground
(?s,SoftwareTestingAndQAengineering) -> Student(?s),
hasFutureSpecialization(?s, SoftwareTester).
```

An example of a SPARQL-query to the ontology is presented below in Fig. 4. This SPARQL-query selects all courses that the student is required to study and incorporates to become a "Manual Tester". Query string:

```
SELECT DISTINCT ?cb WHERE
{?cb rdf:type :SoftwareTester.
:ManualTester :hasCourse ?c }.
```

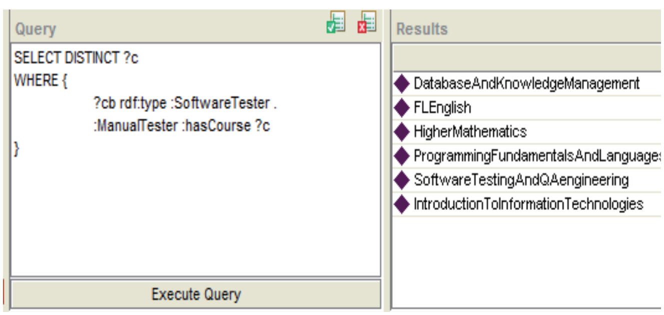


Figure 4. SPARQL-query provided to the ontology

E. Reasoning and decision support in intelligent LMS

The designed prototype of LMS provides the user with an ability to learn courses with the help of user-personalized input. In addition, it provides course recommendations.

The distinctive features of the LMS consist of:

1. Providing recommendations based on the known user data.

2. Providing recommendations based on personal preferences of LMS users.

3. Providing recommendation about future specialization for the users, according to their background knowledge.

4. Providing authorization via a Facebook account.

The target user audience of the system might consist of people who come from a wide variety of age groups and are interested in learning.

The schema of the LMS working process is presented in Fig. 5.

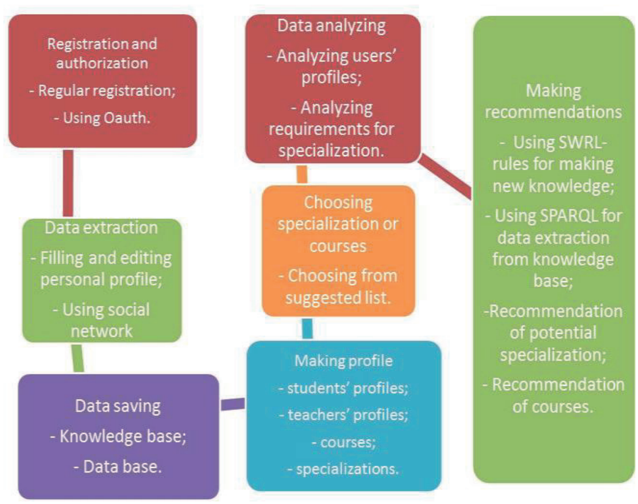


Figure 5. Schema of the LMS working process

In order to log in to the web service, the visitor is asked for a permission to grant access to personal information that is stored in Facebook. Once the user allows access to personal data, he or she is able to view the course's history and get recommendations. This kind of authorization is possible with the help of OAuth technology.

If the user does not have a Facebook account, then he/she can register directly through the user interface of the web service. After login, the user can fill and edit the personal profile. This data will be saved in the knowledge base and in a database for future use. Users have to choose from a suggested list the kinds of courses they have already taken and the kind of specialization or specializations they are interested in.

SWRL-rules created in the knowledge base will add properties like "has future specialization" in users' profiles, according to users' course background. The application, which executes two SPARQL requests, may be developed with the help of a Java-based stack of technologies. One request extracts a list of the required courses from the ontology for further creation of specialization. Another request extracts the student background list from the ontology. After that, the application matches the two extracted lists and shows the courses which the user is required to know to get the desired specialization. As a result, users might get a list of suggested courses.

Conclusions

In this paper, we carried out a literature review about intelligent tools based on semantic technology and about applying these tools in LMS. Based on this research, a model of RS was suggested for LMS to improve the learning process.

An ontology to serve as a knowledge base of LMS users, teachers, courses and specializations was developed. Further recommended specialties for users were formed by SWRL-formulated rules, which allow extraction of new facts from existing knowledge. The Protégé visual environment for OWL development was selected for the implementation of the ontology for the defined subject area, where creation and editing files in RDF, RDFS, OWL, and SWRL-formats and supporting of SPARQL-queries with user-friendly user graphical interface becomes possible. The suggested environment was optimized in accordance with the standards of the World Wide Web Consortium (W3C). The environment is very flexible, and it provides a possibility to quickly modify data and SWRL-rules in the ontology.

The method of extracting personal data from the users' social network profiles with the help of API-functions and the OAuth-protocol was applied. As an example of a social network, we chose Facebook, due to its high social rating by its user adoption factor, its highest leading number of user accounts among other social networks, personal preference, and availability of a documented API that can be used to gain access to users' data.

The architecture of LMS with a decision support tool was created. The method developed takes into account the interests of users while they select their desired specialization and provide some input in choosing courses. It demonstrates an advanced capability of LMSs: implementation of the decision support method for users. The method was also designed for user convenience. It supports users while they choose the most suitable courses as well as increases the popularity of LMSs. In the field of education, the designed model allows making recommendations based on user data extracted from profiles and users' preferences.

The proposed architecture and the model are intended for a development phase to prototype the intelligent tool, where the user's course information is to be entered in a unified input style, and for integration of the proposed intelligent tool into the LMS, for example, Moodle. After that, a number of users will evaluate it: whereas interviews can help to qualitatively assess the outcomes, a survey can help to quantitatively estimate the results to find out if the system architecture was designed suitably, according to user needs.

By continuing the development of this work we can consider and complete the ontology that emerged from a training reason to demonstrate the purpose and opportunity of the technologies involved. It is necessary that many LMSs in future use the same structure of ontology and the same course annotations and user backgrounds to make this system work more efficiently. Later, the system can be compared and evaluated, taking into account not just the studied courses but also the knowledge obtained as an output after the learning courses. Moreover, courses with the same name might have different learning outcomes (for example, one programming course concentrated on Java and another one on Python).

Semantically enhanced decision support can also be used in future for building a career planning or tutoring system for students.

References

[1] Naveh, Gali, Dorit Tubin, and Nava Pliskin. "Student LMS use and satisfaction in academic institutions: The organizational perspective." The Internet and Higher Education 13.3 (2010): 127-133

[2] Meystel, Alexander M., and James Sacra Albus. Intelligent systems: architecture, design, and control. John Wiley & Sons, Inc., 2000.

[3] Schmalzel, John, et al. "An architecture for intelligent systems based on smart sensors." Instrumentation and Measurement, IEEE Transactions on 54.4 (2005): 1612-1616.

[4] Juuso, Esko K. "Integration of intelligent systems in development of smart adaptive systems." International Journal of Approximate Reasoning 35.3 (2004): 307-337.

[5] Santos, O.C.; Boticario, J.G. Educational Recommender Systems and Techniques: Practices and Challenges; IGI Publisher: Hershey, PA, USA; 2011, in press.

[6] Klašnja-Milićević, Aleksandra, et al. "E-Learning personalization based on hybrid recommendation strategy and learning style identification." Computers & Education 56.3 (2011): 885-899.

[7] Zaíane, Osmar R. "Building a recommender agent for e-learning systems."Computers in Education, 2002. Proceedings. International Conference on. IEEE, 2002.

[8] Bobadilla, J. E. S. U. S., Francisco Serradilla, and Antonio Hernando. "Collaborative filtering adapted to recommender systems of e-learning."Knowledge-Based Systems 22.4 (2009): 261-265.

[9] Gutierrez Gallardo, Claudio Domingo, Carlos Hurtado Larraín, and Alberto Mendelzon. "Formal aspects of querying RDF databases." (2014).

[10] Hitzler, Pascal, Markus Krotzsch, and Sebastian Rudolph. Foundations of semantic web technologies. CRC Press, 2011.

[11] Davies, John, Rudi Studer, and Paul Warren, eds. Semantic Web technologies: trends and research in ontology-based systems. John Wiley & Sons, 2006.

[12] Rashid, Sharmin, Ridgewan Khan, and Faysal Ahmed. "A Proposed Model of E-Learning Management System Using Semantic Web Technology." (2013).

[13] Pahl, Claus, and Edmond Holohan. "Applications of semantic web technology to support learning content development." Interdisciplinary Journal of E-Learning and Learning Objects 5 (2009).

[14] Gruber, Thomas R. "Toward principles for the design of ontologies used for knowledge sharing?." International journal of human-computer studies 43.5 (1995): 907-928.

[15] Ellis, R., Calvo, R. A. (2007). Minimum indicators to quality assure blended learning supported by learning management systems, Journal of Educational Technology and Society.

[16] Ellis R. K. (2009). Field guide to learning management systems //Alexandria, VA: ASTD Learning Circuits.

[17] Aydin, C. C., & Tirkes, G. (2010, April). Open source learning management systems in e-learning and Moodle. In Education Engineering (EDUCON), 2010 IEEE (pp. 593-600). IEEE.

[18] Coates, Hamish, Richard James, and Gabrielle Baldwin. "A critical examination of the effects of learning management systems on university teaching and learning." Tertiary education and management 11 (2005): 19-36.

[19] Harper, B., Agostinho, S., Bennett, S., Lukasiak, J., & Lockyer, L. (2005, July). Constructing high quality learning environments using learning designs and learning objects. In Advanced Learning Technologies, 2005. ICALT 2005. Fifth IEEE International Conference on (pp. 266-270). IEEE.

[20] Da Silva, C. B. (2012, June). Pedagogical Model Based on Semantic Web Rule Language. In Computational Science and Its Applications (ICCSA), 2012 12th International Conference on (pp. 125-129). IEEE.

[21] Šimić, Goran, Dragan Gašević, and Vladan Devedžić. "Semantic web and intelligent learning management systems." Workshop on Applications of Semantic Web Technologies for e-Learning. 2004.

[22] Dochev, Danail, and Gennady Agre. "Towards Semantic Web Enhanced Learning." KMIS. 2009.

[23] Kohlhase, Andrea, and Michael Kohlhase. "Semantic knowledge management for education." Proceedings of the IEEE 96.6 (2008): 970-989.

[24] Huang, Weihong, et al. "An intelligent semantic e-learning framework using context-aware Semantic Web technologies." British Journal of Educational Technology 37.3 (2006): 351-373.

[25] Beydoun, Ghassan. "Formal concept analysis for an e-learning semantic web."Expert Systems with Applications 36.8 (2009): 10952-10961

[26] Gladun, Anatoly, et al. "An application of intelligent techniques and semantic web technologies in e-learning environments." Expert Systems with Applications36.2 (2009): 1922-1931

[27] Santos, Olga C., and Jesus G. Boticario. "Requirements for semantic educational recommender systems in formal e-learning scenarios." Algorithms4.2 (2011): 131-154.

[28] Begam, M. Farida, and Gopinath Ganapathy. "Adaptive Learning Management System." International Journal on Soft Computing (IJSC) Vol.4, No.1, 2013

[29] Sivaraman, K. "Effective Web Based E-Learning." Middle-East Journal of Scientific Research 19.8 (2014): 1024-1027.

[30] Pahl, Claus, and Edmond Holohan. "Applications of semantic web technology to support learning content development." Interdisciplinary Journal of E-Learning and Learning Objects 5 (2009).

[31] Dhuria, Shabina, and Sonal Chawla. "Ontologies for Personalized E-Learning in the Semantic Web."

[32] Yankovskaya, Anna E., et al. "Using Semantic Web technologies to create a learning management system based upon Mixed Diagnostic Tests."(2014)

[33] Chi, Y. L. (2009). Ontology-based curriculum content sequencing system with semantic rules. Expert Systems with Applications, 36(4), 7838-7847.

[34] The Protégé Ontology Editor [Online] Available: http://protege.stanford.edu/.

[35] W3C. SWRL: A Semantic Web Rule Language Combining OWL and RuleML [Online] Available: http://www.w3.org/Submission/SWRL/.

PII

ADAPTIVE SYSTEMS AS ENABLERS OF FEEDBACK IN ENGLISH LANGUAGE LEARNING GAME-BASED ENVIRONMENTS

by

Mariia Gavriushenko, Laura Karilainen, Marja Kankaanranta 2015

2015 IEEE Frontiers in Education Conference (FIE), El Paso, Texas, USA

PIII

ADAPTIVE VOCABULARY LEARNING ENVIRONMENT FOR LATE TALKERS

by

Mariia Gavriushenko, Oleksiy Khriyenko, Iida Porokuokka 2016

8th International Conference on Computer Supported Education (CSEDU 2016),
Rome, Italy

Adaptive Vocabulary Learning Environment for Late Talkers

Mariia Gavriushenko, Oleksiy Khriyenko and Iida Porokuokka
Department of of Mathematical Information Technology, University of Jyväskylä, Jyväskylä, Finland
magavriu@student.jyu.fi, {oleksiy.khriyenko, iida.porokuokka }@jyu.fi

Keywords: Adaptive, Self-Adaptive, Vocabulary Learning, Late Talkers, SLI, Game-based Learning.

Abstract: The main aim of this research is to provide children who have an early language delay with an adaptive way to train their vocabulary taking into account individuality of the learner. The suggested system is a mobile game-based learning environment which provides simple tasks where the learner chooses a picture that corresponds to a played back sound from multiple pictures presented on the screen. Our basic assumption is that the more similar the concepts (in our case, words) are, the harder the recognition task is. The system chooses the pictures to be presented on the screen by calculating the distances between the concepts in different dimensions. The distances are considered to consist of semantic, visual and auditory similarities. Each similarity factor can be measured with different methods. According to the user's feedback, the weights of the factors and similarity distance are adjusted to modify the level of difficulty in further iterations. The system is designed to attempt to retrieve knowledge about the learners by recognition of aspects that are difficult for them. Proposed solution could be considered as a self-adaptive system, which is trying to recognize individual model of the learner and apply it for further facilitation of his/her learning process. The use of the system will be demonstrated in future work.

1 INTRODUCTION

The focus of our work is in the area of facilitation learning techniques for children with an early language delay. In the literature, these children are often called late talkers, a term also adopted in this paper. Late talkers are a group of children who learn to form sentences later and have smaller vocabularies than their more typical peers who start putting words together before turning two (Preston et al., 2010). This early language delay has been connected with a risk of later difficulties in language learning such as dyslexia (Lyytinen et al., 2001; Lyytinen, 2015; Lyytinen et al., 2005). A very recent finding (Lyytinen, 2015) confirmed earlier observation (Lyytinen et al., 2005) reveals that if such an early language delay comprises receptive language, i.e. comprehension of spoken language during early years, such children will face serious difficulties in becoming fully literate. Authors documented (Lyytinen et al., 2005) the fact which has been more recently noted also by Nematzadeh et al. (2011), that many late talkers catch up with their peers, but some continue with slower learning pace and are later considered to have a Specific Language Impairment (SLI) (Thal et al., 1997; Desmarais et al., 2008).

Every child is different and learns with different paces and strategies. Information technology has long been seen as a cost-efficient solution for meeting the students' individual needs in learning (Murray and Pérez, 2015). Learning with mobile devices (M-learning) has been recognized to motivate the children to learn as well as to attract their attention while solving problems (Skiada et al., 2014). M-learning can provide a stress-free environment combining ubiquitous learning with individualisation so that the learner gets to proceed in their own pace. As well as games are designed to generate a positive effect in players and are most successful and engaging when they facilitate the flow experience (Kiili, 2005). Flow is considered to be a state of mind where a person forgets his surroundings and track of time while occupying themselves with tasks that are neither too easy (which would lead to boredom), nor too difficult, (which would lead to anxiousness) (Csikszentmihalyi, 1990). In order to facilitate a flow experience meaning to provide sufficient challenge, a learning game should take into account learner's individual needs.

The key to prevent the late talking children from continuing with slow learning pace is to intervene with their vocabulary learning as early as possible. Broadening the vocabulary creates new opportunities to form sentences. As young children are the target group of our study, we faced a challenge of providing them with a simple and motivating way to learn.

In this research, we decided to elaborate an adaptive mobile application with gamified elements such as rewards and animations with actions which are attractive for children. The application is developed for touchscreens that are 7 inches or larger to prevent the effect of motoric skills. The simple functionality on the surface confirms that the system is not too complicated even for very young children.

This paper consists of 4 sections. Section 2 discusses the related work and Section 3 concentrates on late talkers and on how they could benefit from adaptive learning technologies. In Section 4, we describe the proposed system's architecture. In Section 5, conclusions are drawn. Section 6 presents future work and plans for evaluation.

2 RELATED WORK

Many have recognized the importance and the benefits of developing digital systems for young children. It is possible to find lots of game-based learning solutions for preschool children with categories "Preschool games", "Language Arts from Phonics through Reading", "Word Games", "Animal Games", "Phonetics", etc. Unfortunately, they are not made keeping in mind those who have difficulties in learning. Research and development of ICT-supported learning for children with disabilities has not received as much attention and it is also difficult to access research findings in this field (Istenic, Starcic and Bagon, 2014). However, some contributions have been made and in the next paragraphs we discuss few of them.

PAL system proposed by Newell, Booth and Beattie (1991) was created for the children with poor motor control; the key-saving aspects speeded up text creation and it was found very useful for children with spelling problems. Also, it was observed that children who were on the verge of being classified as non-readers, showed a significant improvement in their work.

Skiada et al. (2014) suggested a mobile application named "EasyLexia" which is built taking into account the needs of dyslexic children as well as the usability. EasyLexia included tasks that train reading skills, memory, concentration and mathematical logic providing both auditory and visual stimuli. The results of Skiada et al. showed that children preferred doing the exercises with mobile application over the pen and paper version. Concentrating on the touch screen kept the children's focus better, which emphasizes the use of M-learning as a significant method in today's learning. A fun mobile learning application developed for children with dyslexia certificates the potential of mobile learning application in such environments (Saleh and Alias, 2012). In addition, "Dyseggia" – a game application with word exercises for children with dyslexia suggested in (Rello et al., 2012) showed positive results of the technology use in today's learning methods.

Singleton and Simmons (2001) proposed the multisensory drill-and-practice computer program Wordshark, which is designed to improve children's spelling and word recognition skills. Wordshark is presented in a game-format, and it is used to practice words, learn new words, find out whether children can read and spell particular words. As a reward for good work from these tasks, Wordshark enables earning teaching points. This utility does not have components for assessment or diagnosis and does not automatically adjust the task to the individual learner, but motivates and is a useful reinforcement resource for pupils with dyslexia and also for others with special educational needs.

Possibly the most extensive empirical documentation concerning the efficiency of game based training of the reading skill is coming from the research based on the Graphogame (see graphogame.info). Its effects on the reading skill among children with dyslexia and also among typical learners with insufficient reading instruction in developing countries have been documented in detail in tens of studies listed in the mentioned pages.

While M-learning, as well as game-based learning have proven their benefits in helping children with learning difficulties, there is still not much done, especially in the area of adaptive learning, for the target group of late talkers in vocabulary learning.

3 LATE TALKERS

The child's vocabulary in preschool age is dependent on the social conditions of education. Children who start school with greater literacy skills and background knowledge have a persisting advantage over those children who do not have these skills (Snow et al., 1998).

The connection between an early language delay and later difficulties in language learning has been studied extensively. Preston et al. (2010) mention that longitudinal studies (Scarborough and Dobrich, 1990; Paul et al., 1997; Stothard et al., 1998; Rescorla, 2002, 2005, 2009) have shown that the delay predicts later difficulties in, e.g. reading. Longitudinal study of dyslexia in (Lyytinen et al., 2001) recognized that late talkers with a familial risk for dyslexia are more likely to have such problems than typical children with the same familial risk.

3.1 How Late Talkers Learn

Many studies have noted that late talkers are using different strategies in learning. One of these differences is that some of these children have difficulties in their general cognitive abilities such as attention, categorization and memory skills (Nemanzadeh et al. 2011). Another is related to the connectedness of their semantic network structures which is intuitively connected with the ability of forming sentences. Beckage et al. (2010) and Nemanzadeh et al. (2011) noted that the semantic network structures of late talkers are less connective compared to their age peers. Beckage et al. (2010) retrieved the vocabularies from questionnaires filled by the children's parents and formed the connections between using co-occurrence in a corpus of child directed speech. They also noted that there was more variance in late talkers' network structures than the typical talkers'. Nemantzadeh et al. (2011) studied the matter by teaching novel words to the children, and then connected learned words by the similarity of their meanings – the late talkers learned fewer words and those that were semantically rather further than closer.

The results allow us to conclude that these children require more personalisation during their learning process. The personalisation should allow them to learn with their own pace and keep their attention. In addition, it can support them in forming the associative connections in their vocabulary in order to create better understanding of categorisation.

3.2 Adaptive Learning Systems for Children with Learning Difficulties

It is known that learning is improved when the instructions are given to the learner in a personalised manner (Murray and Pérez, 2015). This knowledge and background theories of education have been a decades lasting trend in creating technologically enhanced learning environments that adapt, one way or another, to the learners needs. In the literature, these learning environments are often referred as "adaptive" or "intelligent" tutoring or learning systems.

According to (Gifford, 2013), adaptive learning is a methodology that is centred on "creating a learning experience that is unique" for every individual learner through the intervention of computer software. Adaptive learning systems allow organising content, identifying the way to learn according to learner's knowledge and use assessment result to provide personalised feedback for each learner (Sonwalkar, 2005).

Adaptive learning systems have a lot of features and functions (Venable, 2011) that are combined to provide relevant content, support and to guide the user through the adaptive learning courses or modules: pre-test, pacing and control, feedback and assessment, progress tracking and reports, motivation and reward.

These systems can be either simple or algorithm-based (Oxman et al., 2014). Simpler adaptive learning systems are rule-based, created using a series of if-then statements. Algorithm-based systems take advantage of advanced mathematical formulas and machine learning concepts to adapt with greater specificity to individual learners. Earlier research by (Brusilovsky and Peylo, 2003) divides these systems into Adaptive Hypermedia Systems and Intelligent Tutoring Systems. By these technologies we mean different ways to add adaptive or intelligent functions into learning systems. Adaptive Hypermedia Systems include adaptive presentation and adaptive navigation support, and also adaptive information filtering, which includes collaborative filtering and content-based filtering. Intelligent Tutoring Systems include curriculum sequencing, intelligent solution analysis and problem solving support and intelligent collaborative learning, which includes adaptive group formation and adaptive collaboration support.

Finding an optimal way to present the concepts to the children during their first years of life might benefit in diminishing their risk to develop difficulties in language learning in the future. That is why we should have adaptive content presentation in learning system for late talkers. To make it available we should use the adaptive presentation technology, which aims to adapt the content of a hypermedia page to the user's goals, knowledge and other information stored in the user model. In a system with adaptive presentation, the pages are not static, but adaptively generated or assembled from pieces for each user (Brusilovsky, 1999). Respectively, in our system, the content shown to the children should not be predefined, but should have the ability to adapt according to user's feedback.

We assume that it is harder to recognize some particular concept among several other similar ones. Every child can perceive images and sounds corresponding to these images differently. For one child, it is difficult to distinguish between words, which sound similar. For another child, it could be difficult to distinguish between words that have a similar pattern, whether they could have a similar shape or similar colour. Another child may have difficulty with differentiating words, which are semantically close to each other.

With proposed solution we develop an adaptive learning system for late talkers, which will take into account personal qualities of learners' in their perception of a concept. The system will learn and build a personal model of a learner based on his/her answers while changing the complexity of concept representation, and will apply this model for further facilitation of learning process. Therefore, the system could speed up the vocabulary learning process individually for each user.

4 SYSTEM DESCRIPTION

As a basis for the learning system improvement, we have chosen "Graphogame" learning tool (Richardson and Lyytinen, 2014), which is a learning environment that teaches children reading by playing back a sound of a letter and asking the learner to choose the correct letter from multiple choices presented on a screen, adapting the given choices according to similarities between the letters and user's feedback. We are going to utilise the functionality of "Graphogame" tool replacing the letters with images of words. We are not going to use predefined sets of images or populate a set with random images because such strategy lacks an individual approach and does not take personal specifics of a learner into account. Therefore, the functionality of our approach should be able to take into account differences between visual and auditory stimuli and personalise further picture selection based on intelligent analysis of user's answers.

4.1 Concept Similarity

We make an assumption that concept learning and recognition depends on individual perception of several parameters such as sound and visual representations as well as their semantics. Our approach is based on manipulation with complexity level of concept recognition caused by these factors. Taking into account users' feedbacks (correctness of answers), system will automatically increase or decrease the level of complexity, adapting it to the individual learning abilities of the users, and in such a way will facilitate learning process.

We highlighted three main factors that influence the complexity. These factors are: visual similarity, similarity of sound-based phonetic representations and semantic similarity of the concepts. Assuming that sets of more similar concepts bring more difficulties for their recognition (distinction), we automate the process of image selection using multidimensional concept similarity metric, defined

$$ConceptSim\left(c_i, c_j\right) = F(Sim_V, Sim_P, Sim_S) \qquad (1)$$

as an aggregation of similarity values of mentioned three factors. Now, we are able to personalize our learning tool via changing the delta (Δ) of concept similarity as well as "fiddling" with different levels of influence of the three factors on aggregated similarity.

4.1.1 Visual Similarity

Measuring the distance between two images is a central problem in image recognition and computer vision and many definitions for the metric have been suggested (Wang et al., 2005). Considering visual similarity, image features such as shape of presented object, colour distribution, brightness, contrast, etc. are usually acknowledged. The most commonly used image metric is Euclidean distance due to its advantage of simplicity. Euclidean distance is computed by summing the squares of differences between each pixel in images (Wang et al., 2005). However, in pattern recognition, it performs poorly compared to e.g. Tangent distance (Simard et al., 1993) which succeeds well in tasks such as recognizing handwritten digits (Wang et al., 2005). On shape matching tasks, e.g. Latecki and Lakämper's, (2000) approach gives intuitive results basing the metric on correspondence between object's visual parts. Image Euclidean Distance Measure (IMED) metric (Wang et al., 2005) is based on Euclidean distance. IMED adds the spatial relationships of pictures into consideration and outperformed traditional Euclidean distance in face recognition tasks in evaluations performed by the authors.

Besides shape, colour is a very dominant visual feature. According to (Deng et al., 2001), distance between colours in two pictures can be measured by comparing their colour histograms, which represent the colour distribution in an image. However, colour histograms do not consider spatial knowledge and have high cost in retrieval and search. Authors proposed a "dominant colour descriptor" which consists of the representative colours in a region and their distribution. A similarity measure for the descriptor and an efficient colour indexing scheme for image retrieval were also suggested. The method performed fast and efficiently in the experimentations. However, the descriptor did not take into account the spatial relationships between the colour regions and considering high level matches, the correspondence was unstable.

Some of the visual similarity features might be more valuable than others. For example, take a set of images created by the same designer: they might be drawn in the same style using the same colour palette that makes all images quite similar in spite of the difference between the objects they represent. In this case, shape feature might be more valuable in calculating the actual human perception of visual similarity. Thus, there might be various automatic techniques to recognize different levels of the features relevance, but in our solution, we are going to use manually defined coefficient for the feature and leave possible automation of this process for future work.

4.1.2 Phonetic Similarity

Along with visual similarity, phonetic similarity of concepts can also affect children's performance in distinguishing them. There are many researches and practical implementations done with respect to automated voice and speech recognition (Petajan, 1990; Astradabadi, 1998; Potamianos et al., 1997) Adaptive speech recognition technology is not yet at the point where machines understand all speech, in any acoustic environment, or by any person, but it is used on a day-to-day basis in a number of applications and services (Docsoft Inc., 2009). The vast difference in anatomy and physiology between the speech production and perception systems of humans makes it difficult to analyse (Kessler, 2005). Unfortunately, these systems are expensive and they cannot always correctly recognize the input from a person who speaks with a dialect, accent, and also they have some problems with recognizing words from people who are combining words from different languages by force of habit.

Because the complexity of voice recognition especially in case of sound samples created by different persons is considered high, we decided to calculate phonetic similarity of the concepts as string based similarity of their phonetic transcriptions. In a non-orthographic language, the phonetic representations of the words are more reliable than the written format of words. However, they do not take into account e.g. different dialects that would sound more native to the learner. In spite of these disadvantages, it was decided that the standardised language provides enough information on the phonetic distances. The auditory representation of the words is provided in such a way that it follows the patterns of these phonetic representations aiming to the most standardised way of speaking.

There are several string-based techniques that could be applied for phonetic transcriptions similarity matching: Edit Distance – finds how dissimilar two strings are by counting the minimum number of operations required to transform one string into another; Jaro-Winkler measure (Winkler, 1999), N-gram similarity function (Kondrak, 2005), Soundex (Russell and Odell, 1918) – phonetic similarity measure, which principle of operation is based on the partition of consonants in the group with serial numbers from which then compiled the resulting value; Daitch-Mokotoff (Mokotoff, 1997) has much more complex conversion rules than in Soundex – now shaping the resulting code involved not only single characters, but also a sequence of several characters; Metaphone – transforms the original word with the rules of English language, using much more complex rules, and thus lost significantly less information as letters are not divided into groups (Euzenat and Shvaiko, 2013). In our solution we allow utilisation of several measuring functions with further weighted aggregation of the results (e.g., weighted product or weighted sum).

4.1.3 Semantic Similarity

Semantic similarity between concepts plays an important role in semantic sense understanding. The measuring is not a trivial task. The most used technique is to measure semantic similarity based on domain ontology; a conceptual model that describes the corresponding domain. Ontology-based semantic similarity can be measured with different methods (Sanchez, 2012; Bin et al., 2009):

- Edge-counting method (or graph-based) calculates the minimum path length connecting the corresponding ontological nodes through the 'is-a' -links (Sanchez et al., 2012). Equally distant pairs of concepts which belong to the upper level of taxonomy are counted less similar than those which belong to a lower level (Wu and Palmer, 1994). Very often the shortest path length is combined with the depth of ontology in a nonlinear function (Li et al., 2003), sometimes with the overlapping between the nodes (Alvarez and Lim, 2007). In (Al-Mubaid and Nguyen, 2006), authors applied cluster-based measure on top of a minimum length path and taxonomical depth. Taxonomical edge-counting method was extended by including non-taxonomic semantic links to the notion in the path (Hirst and St-Onge, 1998).
- Feature-based measures use taxonomical features extracted from ontology. The similarity between two concepts can be computed as a function of their common and differential features (assessing similarity between concepts as a function of their properties) (Sanchez et al., 2012). Such facet-based classification could be combined with similarity of common properties' values.
- Combined measures which include the edge-counting based and information content (IC)-based measures with edge weight (Jiang and Conrath, 1997).

In our current solution, we have limited our focus to a domain of animals. Each domain brings certain specifics to semantic similarity measuring metric. Semantic similarity could be measured differently depending on the context. Such context dependent similarities could be calculated separately and be further aggregated using different weights for different contexts. Thus, for the chosen domain, we may highlight several classifications:

- Biological species-based classification: this metric is based on subclass hierarchy of animals classified by biological families of animals. In this case, the most suitable approaches to measure similarity are graph-based techniques (e.g. Jaccard metric, Scaled shortest path, Depth of the subsumer and closeness to the concepts, etc.) (Euzenat and Shvaiko, 2013; Bouquet et al., 2004; Leacock and Chodorow, 1998; Haase e. al., 2004). Thus, we calculate semantic similarity of concepts based on locations of corresponding nodes in the graph, using taxonomy-based ontology that represents class hierarchy of animals.
- Geolocation-based classification: here we distinguished animals by geographical regions they live in. It is a complex metric that integrates continent-based, latitude- and climate zone based clustering. Similarity between the clusters is calculated based on climate groups' hierarchy and similarity of different continents.
- Domestication-based classification: animals could be also divided to those who are fully domesticated by human and live at their homes and farms, those whom we may meet in a zoo, and those who live only in wild nature and most probably are only seen via video records and photos. In this case, distances between the classes could be predefined. Also, other metrics that define semantic similarity in other contexts may exist. Therefore, analogically to other similarity factors, final semantic similarity measure could be aggregated by weighted products/sum.

4.1.4 Concept Similarity Measure

Since all of our concept similarity factors (visual, phonetic and semantic) could be represented as weighted functions of various similarity measuring techniques, we may define a general formula to calculate similarity between the concepts (Figure 1).

4.2 System Architecture and Adaptation Logic

The system is presented as game-based learning environment for mobile phones and tablets. The system's aim is to teach recognition of vocabulary items from multiple choices of their visual and auditory representations. The task of the learner is to recognize the word that he or she hears from a group of pictures presented on the screen. All pictures except for the one presenting the correct answer are further referred as distraction items. General architecture of proposed facilitation solution is shown in Figure 2.

$$ConceptSim\,(c_i, c_j) = w_V * \sum_{k_V=0}^{n_V} w_{k_V} * Sim_{V_{k_V}} + w_P * \sum_{k_P=0}^{n_P} w_{k_P} * Sim_{P_{k_P}} + w_S * \sum_{k_S=0}^{n_S} w_{k_S} * Sim_{S_{k_S}}$$

where w_V, w_P, w_S *are coefficients of weighted influence of visual, phonetical and semantic factors on concept similarity;* n_V, n_P, n_S *are number of different similarity measuring techniques used to calculate visual (*Sim_V*), phonetical (*Sim_P*) and semantic (*Sim_S*) similarity with corresponding weights of their influence (*$w_{k_V}, w_{k_P}, w_{k_S}$*).*

Figure 1: Concept similarity measuring function.

Based on calculated similarity measures between all the concepts and the chosen one ("Random Concept" in the figure), we rank them and select a group of the most closest to the defined delta (Δ) value. Initial value could be, for instance, considered as an average of similarity values of all the measured concepts. Depending on the amount of distraction items N_c (since their amount also influences the overall complexity), system chooses the concepts with similarity value closest to the value of Δ. Depending on the user's feedback (answer), the value of Δ will be changed in the feedback analysis module. If a user makes a mistake and provides wrong answers, the Δ value will be increased (moved to the side of the concepts with lower similarity). Otherwise, complexity could be increased (by decreasing value, a group of more similar concepts will be selected next time).

At the same time, manipulating with coefficients of visual, phonetic and semantic factors influence (the vector of weights, $\vec{W} = (w_V, w_P, w_S)$ we are able to recognize levels of difficulties that an individual factor brings for a particular learner. Once user provides a wrong answer, next time, when the same concept will be chosen, the algorithm will change vector of weights giving more preference to one of the factors. Therefore, system will collect statistics on personal learning model of the user, while trying to already personalize complexity of the tasks. Whole collected statistics including value of Δ, vector of weights $\vec{W}$, values of similarities between chosen and other distraction concepts, etc. is stored in the Log module of the system and is further used as labelled learning sample to recognize individual user's features of concept perception. Furthermore, this will allow the system to personalize the strategy for individual learning process and to develop ability of the learner to overcome individual difficulties in vocabulary learning.

For the current research we used breadth-first search to calculate the semantic similarity from the graph of used concepts, Levenhstein distance for phonetic similarity and Euclidean distance for the visual similarity. In future work, we will use other algorithms, as well as combine and compare them to conclude which are the most relevant.

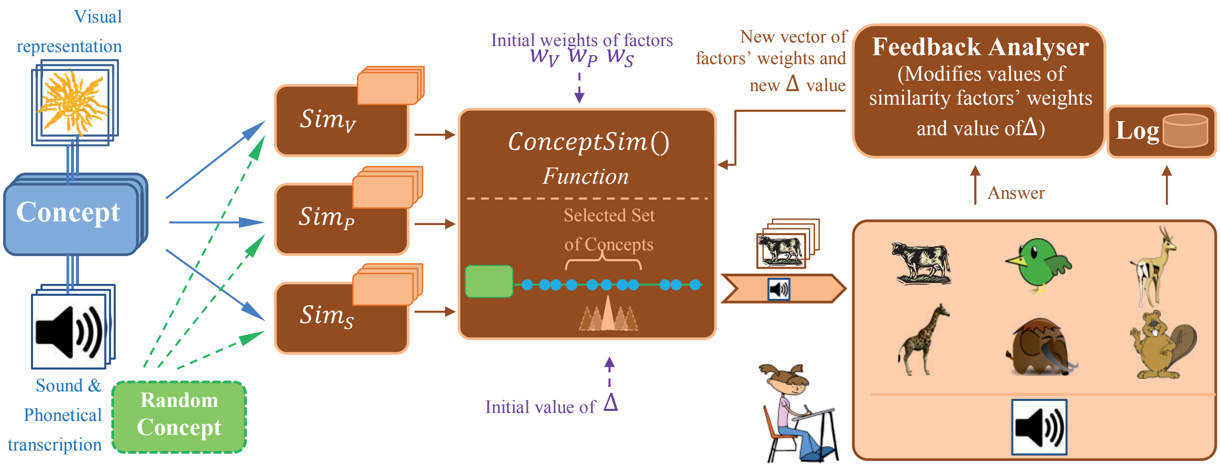


Figure 2: General architecture of solution.

5 CONCLUSIONS

In this work, we concentrated on developing an adaptive vocabulary learning system for late talkers who have difficulties in learning language due to individual reasons. The system is a mobile learning application which aims to provide an optimal way of vocabulary presentation for young children.

The functionality behind the system is based on different similarity factors (visual, phonetic and semantic) of the learning objectives, words, in our case. We assumed that the more similar the words are, the more difficult it is to distinguish them. Thus, as the learners improve, the system is able to provide them with more challenging tasks. In addition to using the information of learners' mistakes in the system's adaptation, the data is collected and further analysed from angles of semantic network connectivity, emphasis of the dimensions and overall system's educational effectivity.

Existing approaches to study the characteristics of semantic networks are dependent on either parent's knowledge of their children's vocabularies (cf. Beckage et al, 2010). Our approach has the advantage of not only impartiality but automaticity, which makes it possible to collect larger amount of data with less effort and bias. However, some disadvantages in the methodology exist. Firstly, the system does not require the children to form the words, only to recognize them, and therefore it cannot be considered as a way of training word composition. Secondly, it is likely that all the children's answers are not equally valuable since multiple-choices questions can be answered randomly and also wrong answers could be given intentionally.

6 FUTURE WORK

To evaluate the system's efficiency, it would be tested with a group of children consisting of both 1-3-year-old late talkers and other children with more typical language proficiency. The testing group of children will be split in halves. For every child a filtration process of the concepts will be made. By testing, the system will exclude the words that child already knows from the provided sample of the animal words and will only work with those words child does not know. For the first group of children, the minimum delta (lowest difficulty), which is not changed during the learning and testing process, would be used. And for the second group, the delta will be changing for every learner personally according to the adaptation logic of the system. After some period of learning and testing there will be a final test, which would check the amount of words the children learned. Knowing the percentage of the learned concepts of each group, we can make a conclusion if the proposed adaptive system is effective. After the evaluation we will use the gathered data for further development of the algorithm. As the current system relies on metrics made by adults, it is possible that they do not reflect the children's associations. The gathered association networks can be further used as a base for the algorithm instead of the adult-made ontologies. The evaluation results could also show if recognizable patterns exist in the mistakes the children make.

Also, we will extend the knowledge base, including also other domains of vocabulary. The data like user's answers could be automatically sent to a database on a cloud server. The algorithm could be extended to consider several users' feedback, which could be used to form more reliable user models. Applied to a larger set and multiple vocabulary domains, analysis of the children's answers in the system can result in valuable information of children's semantic networks.

The system could also be modified to be more intuitive and motivating. For example, it could be extended to provide interactive storybook-type of tasks, where a child could listen to a story illustrated on the screen and then answers to multiple-choice questions about the context of the story.

ACKNOWLEDGEMENTS

We sincerely thank Professor Heikki Lyytinen for the inspiration, guidance and encouragement in finishing this part of work, as well as helping in future development.

REFERENCES

Al-Mubaid, H., Nguyen, H. A., 2006. A cluster-based approach for semantic similarity in the biomedical domain. In *28th Annual International Conference of the IEEE Engineering in Medicine and Biology* New York, USA: IEEE Computer Society.

Alvarez, M. A., Lim, S., 2007. A Graph Modelling of Semantic Similarity between Words, In *Proc. 2007*

International Conference on Semantic Computing, pp. 355- 362.

Astarabadi, S., 1998. Automated retrieval of voicemail using speech recognition. U.S. Patent No. 5,822,405. 13 Oct. 1998.

Beckage, N., Smith, L., Hills, T., 2010. Semantic network connectivity is related to vocabulary growth rate in children. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, Portland, OR, USA (pp. 2769-2774).

Beckage N., Smith L., Hills T., 2011. Small Worlds and Semantic Network Growth in Typical and Late Talkers. PLoS ONE 6(5): e19348. doi:10.1371/journal.pone.0019348.

Bin, Shi, et al. "Ontology-based measure of semantic similarity between concepts." *Software Engineering, 2009. WCSE'09. WRI World Congress on*. Vol. 2. IEEE, 2009.

Bouquet, P., Kuper, G., Scoz, M, Zanobini, S., 2004. Asking and answering semantic queries. In *Proc. of Meaning Coordination and Negotiation Workshop (MCNW-04) in conjunction with International Semantic Web Conference (ISWC-04)*, 2004.

Brusilovsky, P., 1999. Adaptive and intelligent technologies for web-based eduction. *KI* 13.4: 19-25.

Brusilovsky, P., Peylo, C., 2003. Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education* 13.2: 159-172.

Csikszentmihalyi, M., 1990. Flow: The Psychology of Optimal Experience. *Harper Perennial,* New York.

Deng, Y., Manjunath, B.S., Kenney, C., Moore, M.S., Shin, H., 2001. An efficient color representation for image retrieval, in *Image Processing, IEEE Transactions on , vol.10, no.1,* pp.140-147, Jan, doi: 10.1109/83.892450.

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., Rouleau, N., 2008. Systematic review of the literature on characteristics of late-talking toddlers. *Int'l J. of Lang. and Communication Disorders*, 43(4), 361–389.

Docsoft inc., 2009. What is Automatic Speech Recognition? retrieved from: http://www.docsoft.com/resources/Studies/Whitepapers/whitepaper-ASR.pdf.

Euzenat, J., Shvaiko, P., 2013. Ontology matching. Springer-Verlag, *Berlin Heidelberg (DE)*, 2013, isbn 978-3-642-38720-3.

Gifford T., 2013. Take It Personally: Adaptive Learning. retrieved from http://eltjam.com/adaptive-learning/

Haase, P., Siebes, R., van Harmelen, F., 2004. Peer Selection in Peer-to- Peer Networks with Semantic Topologies. In *Proc. of International Conference on Semantics of a Networked World: Semantics for Grid Databases*.

Hills, TT, Maouene, J, Riordan, B, Smith, LB, 2010. The Associative Structure of Language: Contextual Diversity in Early Word Learning. *Journal of memory and language 63*: 259–273.

Hills, TT, Maouene, M, Maouene, J, Sheya, A, Smith, LB, 2009. Longitudinal analysis of early semantic networks: preferential attachment or preferential acquisition? *Psychological Science 20*: 729–739.

Hirst, G., & St-Onge, D. (1998). Lexical chains as representations of context for the detection and correction of malapropisms. In C. Fellbaum (Ed.), WordNet: An Electronic Lexical Database (pp. 305–332). MIT Press.

Istenic Starcic, A. and Bagon, S., 2014. ICT- supported learning for inclusion of people with special needs: Review of seven educational technology journals, 1970–2011. *British Journal of Educational Technology*, *45*(2), pp.202-230.

Jiang, J., Conrath, W., 1997. Semantic Similarity Based On Corpus Statistics and Lexical Taxonomy. In *Proceedings of International Conference Research on Computational Linguistics.*

Jones, S., Smith, L., 2005. Object name learning and object perception: a deficit in late talkers. *Journal of Child Language, 32*, pp 223-240. doi:10.1017/S0305000904006646.

Kessler, B., 2005. Phonetic comparison algorithms, *Transactions of the Philological Society Volume 103*:2 243–260.

Kiili, K., 2005. Digital game-based learning: Towards an experiential gaming model. *The Internet and higher education 8.1*: 13-24.

Kondrak, G., 2005. *N-gram similarity and distance. In String processing and information retrieval* Springer Berlin Heidelberg: 115-126.

Latecki, L.J., Lakamper, R., 2000. Shape similarity measure based on correspondence of visual parts. In *Pattern Analysis and Machine Intelligence*, IEEE Transactions on , vol.22, no.10, pp.1185-1190, Oct.

Leacock, C., Chodorow, M., 1998. Combining local context and WordNet similarity for word sense identification, In *Fellbaum MIT Press*, pp. 265–283, 1998.

Li, Y., Bandar, Z., McLean, D., 2003. An approach for measuring semantic similarity between words using multiple information sources. *Knowledge and Data Engineering*, IEEE Transactions on, 15(4), 871-882.

Lyytinen, P, Poikkeus, A-M, Laakso, M-L, Eklund, K, Lyytinen, H., 2001. Language development and symbolic play in children with and without familial risk for dyslexia. *J Speech Lang Hear Res* 2001; 44: 873–85.

Lyytinen, Paula, Kenneth Eklund, and Heikki Lyytinen. "Language development and literacy skills in late-talking toddlers with and without familial risk for dyslexia." *Annals of dyslexia* 55.2 (2005): 166-192.

Mokotoff, G., 1997. Soundexing and genealogy retrieved from http://www.avotaynu.com/soundex.htm.

Murray, Meg Coffin, and Jorge Pérez. 2015. *Informing and Performing: A Study Comparing Adaptive Learning to Traditional Learning*. Informing Science: The International Journal of an Emerging Transdiscipline.

Nematzadeh, A., Afsaneh, F., Stevenson, S., 2011. A computational study of late talking in word-meaning acquisition. *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*.

Newell, A.F., Booth, L. and Beattie, W., 1991. Predictive text entry with PAL and children with learning difficulties. *British Journal of Educational Technology*, *22*(1), pp.23-40.

Oxman, S., Wong, W., D. V. X. Innovations, 2014. White Paper: Adaptive Learning Systems. *Snapwiz. Integrated Education Solutions,* 1.

Paul, R., Murray, C., Clancy, K., Andrews, D., 1997. Reading and metaphonological outcomes in late talkers. *J Speech Lang Hear Res 1997*; 40: 1037–47.

Petajan, E., 1990. Electronic facial tracking and detection system and method and apparatus for automated speech recognition. *U.S. Patent No. 4,975,960*. 4 Dec. 1990.

Potamianos, A., Shrikanth, N., Sungbok, L., 1997. Automatic speech recognition for children. *Eurospeech*. Vol. 97. 1997.

Preston, J., Frost, S., Mencl, W., Fulbright, R., Landi, N., Grigorenko, E., Jacobsen, L., Pugh, R., 2010. Early and late talkers: school-age language, literacy and neurolinguistic differences. *Brain 133 (8):* 2185–2195. ISSN: 0006-8950. doi:10.1093/brain/awq163.

Rello, L., Bayarri, C., Gorriz, A., 2012. What is wrong with this word? Dyseggxia: a game for children with dyslexia. In *Proceedings of the 14th international ACM SIGACCESS conference on Computers and accessibility* (pp. 219-220). ACM.

Rescorla, L., 2002. Language and reading outcomes to age 9 in late-talking toddlers. *J Speech Lang Hear Res* 2002; 45: 360–71.

Rescorla, L., 2005. Age 13 language and reading outcomes in late-talking toddlers. *J Speech Lang Hear Res 2005*; 48: 459–72.

Rescorla, L., 2009. Age 17 language and reading outcomes in late-talking toddlers: Support for a dimensional perspective on language delay. *J Speech Lang Hear Res* 2009; 52: 16–30.

Richardson, U., Lyytinen, H., 2014. The GraphoGame method: the theoretical and methodological background of the technology-enhanced learning environment for learning to read. *An Interdisciplinary Journal on Humans in ICT Environments* 10 (1): 39–60. ISSN: 1795-6889. doi:http://dx.doi.org/10.17011/ht/urn.201405281859.

Russell, R., M. Odell.,1918. Soundex. *US Patent 1* (1918).

Saleh, R., Alias, N.A., 2012. Learner Needs Analysis for Mobile Learning Comic Application among Dyslexic Children*, International Journal Of Education And Information Technologies, vol. 6,* no. 2, 2012.

Sánchez, D., Batet, M., Isern, D., Valls, A., 2012. Ontology-based semantic similarity: A new feature-based approach, *Expert Systems with Applications*, Volume 39, Issue 9, July 2012, Pages 7718-7728, ISSN 0957-4174, http://dx.doi.org/10.1016/j.eswa.2012.01.082.

Scarborough, HS, Dobrich W., 1990. Development of children with early language delay. *J Speech Hear Res* 1990; 33: 70–83.

Simard, P., LeCun, Y., Denker, J. S., 1993. Efficient pattern recognition using a new transformation distance. In *Advances in neural information processing systems* (pp. 50-58).

Singleton, C. and Simmons, F., 2001. An evaluation of Wordshark in the classroom. *British journal of educational technology*, *32*(3), pp.317-330.

Skiada, R., Soroniati, E., Gardeli, A., Zissis, D., 2014. EasyLexia: A mobile application for children with learning difficulties. Procedia *Computer Science*, 27, 218-228.

Snow, C. E., Burns, M. S., & Griffin, P. (Eds.) (1998). Preventing Reading Difficulties in Young Children. Washington, DC: National Academy Press.

Sonwalkar, N., 2005. Adaptive learning technologies: From one-size-fits-all to individualization. *EDUCAUSE Center for Applied Research, Research Bulletin*(2005).

Stothard, SE, Snowling, MJ, Bishop, DVM, Chipchase, BB, Kaplan, CA., 1998. Language-impaired preschoolers: a follow-up into adolescence. *J Speech Lang Hear Res* 1998; 41: 407–18.

Thal, D. J., Bates, E., Goodman, J., Jahn-Samilo, J., 1997. Continuity of language abilities: An exploratory study of late- and early-talking toddlers. *Developmental Neuropsychology, 1*3(3), 239–273.

Venable, M., 2011. Adaptive Learning Technology: An Introduction. Web-site: http://www.onlinecollege.org/2011/08/30/adaptive-learning-technology-an-introduction/

Wang, L., Zhang, Y., Feng, J., 2005. On the Euclidean distance of images. In *Pattern Analysis and Machine Intelligence*, IEEE Transactions on, vol.27, no.8, pp.1334-1339, Aug. 2005.

Winkler, W.E., 1999. The state of record linkage and current research problems. *In Statistical Research Division*, US Census Bureau.

Wu, Z., & Palmer, M. (1994). Verb semantics and lexical selection. In *32nd annual Meeting of the Association for Computational Linguistics* (pp. 133–138). Las Cruces, New Mexico: Association for Computational Linguistics.

PIV

AN INTELLIGENT LEARNING SUPPORT SYSTEM

by

Mariia Gavriushenko, Oleksiy Khriyenko, Ari Tuhkala 2017

9th International Conference on Computer Supported Education (CSEDU 2017),
Porto, Portugal

An Intelligent Learning Support System

Mariia Gavriushenko, Oleksiy Khriyenko and Ari Tuhkala
Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
magavriu@student.jyu.fi, oleksiy.khriyenko@jyu.fi, ari.tuhkala@gmail.com

Keywords: Intelligent learning system, Adaptive and personalized education, Career development.

Abstract: Fast-growing technologies are shaping many aspects of societies. Educational systems, in general, are still rather traditional: learner applies for school or university, chooses the subject, takes the courses, and finally graduates. The problem is that labor markets are constantly changing and the needed professional skills might not match with the curriculum of the educational program. It might be that it is not even possible to learn a combination of desired skills within one educational organization. For example, there are only a few universities that can provide high-quality teaching in several different areas. Therefore, learners may have to study specific modules and units somewhere else, for example, in massive open online courses. A person, who is learning some particular content from outside of the university, could have some knowledge gaps which should be recognized. We argue that it is possible to respond to these challenges with adaptive, intelligent, and personalized learning systems that utilize data analytics, machine learning, and Semantic Web technologies. In this paper, we propose a model for an Intelligent Learning Support System that guides learner during the whole lifecycle using semantic annotation methodology. Semantic annotation of learning materials is done not only on the course level but also at the content level to perform semantic reasoning about the possible learning gaps. Based on this reasoning, the system can recommend extensive learning material.

1 INTRODUCTION

Many young people face difficulties when trying to decide what kind of education to apply. Once the approval to the wanted education is achieved, many decisions must be made: which subjects to choose, what courses to put to the personal study plan, what courses are relevant in future job markets, or if the official curriculum should be supported with supplemental education from the private sector. These decisions might have life-long effects, but often the resources for making the right choices are limited. People should recognize their competitive advantages and put them in the perspective of global trends, technological development, and labor markets. Not only young people have such need, but many graduates and experienced workers also face the situation of economic change, and suddenly they have to adjust their competence by improving and extending their skills.

In welfare countries, flexible education system offers quite a many options to choose. The possibilities are so widespread that it is not easy to make the right choice without a comprehensive analysis of person's capabilities, ambitions, and values against the future labor markets. However, only a small amount of people have an actual freedom to choose their future profession. Usually, people have to be competitive enough to survive, especially in the times of crisis, and those who have the possibility, are changing their location to get a better education for themselves. On the contrary, it has become possible to get the education from another side of the world, especially in higher education. Massive open online courses and learning platforms like Udemy, Khan Academy, and Coursera are helping people to get desired skills (e.g., programming, developer skills, and data analytics) without leaving home. However, some courses have pre-requirements, so that it would be quite difficult to attend courses without filling the existing knowledge gaps by studying an additional learning material.

Educational systems are structured around curricula that aim to supply people with the defined competencies. The problem is that employers are more and more finding people with various skills and backgrounds, which may not be in line with the curriculum. Therefore, people should be able to create specific curriculum and be educated following personal learning trajectory. To meet this demand, topics around different learning support systems have become popular. For example, the Open University of

Hong Kong has developed a system that instantly responds to inquiries about career development, program and course choices, study plans, and graduation checks (Leung et al., 2010). While these systems have potential, the problem is that they should be connected with various educational and career planning services. We can already verify this issue by looking how many different job portals there exists, both nationally and globally.

We argue that there is a need for an Intelligent Learning Support System (ILSS), which will be based on current research information about adaptive, personalized, and intelligent tools. We begin by reviewing the current body of knowledge about these topics. Next, we explain the whole ecosystem of ILSS and the components it includes. Finally, we discuss implications that must be considered when implementing this kind of system. This proposition paper is current because the development of adaptive, personalized, and intelligent systems for education are still in the premature phase in Finland, where technological infrastructure is Finland under reconstruction: The government's aim is to standardize and integrate the social and health care systems, including the career and employment services. Students, job seekers, employment officials, and company headhunters would benefit from a centralized system that aims to make educational interests and employment supply meet.

2 LITERATURE REVIEW

Each person has different knowledge background, ability to memorize, learning speed, motivation, and preferences. Furthermore, people have different health conditions and habits when and how to learn. The problem with educational applications is that they are built for a specific group of learners and not individually. Therefore, development of learning environments with individualized learning mechanisms is an important research topic. Researchers are interested in developing learning environments, which could adaptively provide learning path (Chen, 2008), but the challenge is the development of advanced learning applications that include intelligence and adaptivity (Brusilovsky and Peylo, 2003).

Intelligent learning systems are related to adaptive learning systems and have an intersection (Brusilovsky and Peylo, 2003). Intelligent systems are "systems which use techniques from Artificial Intelligence to provide better support and feedback for learners" (Brusilovsky and Peylo, 2003) and adaptive systems are "systems which attempt to be different for each learner or group of learners" according to information about learners received during their actions (Brusilovsky and Peylo, 2003). Personalized systems present a specific type of general adaptive systems (García-Barrios et al., 2005) and they mean adaptation towards a specific learner.

Computer based training or computer aided learning does not use a model of the learner's knowledge and does not take into account personality of the learner, but it just uses traditional instructional methods (Phobun and Vicheanpanya, 2010). The combination of adaptation, intelligence and personalization could make this better. The interaction with the learner would not only report the correct or incorrect answers, but also explanations of the error cause, and recommendations for study material.

The development of adaptive, intelligent, and personalized learning systems is important, especially for children with disabilities, because they require more personalization during their learning process. These techniques should allow them to learn at their pace and help to keep their motivation (Gavriushenko et al., 2016).

2.1 Personalized learning systems

Many authors suggest in their work different tools for learning systems, which could assist in learning and make this process more personalized. For example, in the work (Bendakir and Aïmeur, 2006) authors presented a course recommendation system which analyzes past behavior of learners concerning their choices. Authors used data mining association rules (user ratings), and this classification makes possible to build a decision tree that classifies each learner profiles. The system recommends learning path by checking which courses are followed simultaneously. Unfortunately, proposed system does not consider learners' background and course availability. Also, researchers (Werghi and Kamoun, 2009) have presented a decision support system for learner advising which provides an automated program planning and scheduling service that fits best their profiles while meeting academic requirements. The system was based on decision tree algorithm. The work (Jeong et al., 2012) presented a personalized learning course planner with decision support using user profile. This system allows the learner to select the learning course they desire taking into account previous learning information of learner. Authors used organization algorithm containing the decision matrix. The proposed system improved learning effectiveness and especially learners' satisfactory according to the questionnaire. Also, in a paper (Jyothi et al., 2012) was presented a recommender system which assists the instructor in building learning path, more specifically in identifying the groups of learners who have similar learning styles and then providing recommendations to these learners. Authors used Felder-Silverman learning style model.

2.2 Intelligent and adaptive learning systems

Some research papers were concentrated on developing intelligent and adaptive applications for learning which could fully or partially reduce the involvement of human advisers. For example, in the work (Henderson and Goodridge, 2015) authors proposed a system which gives accurate advice only to those pursuing "special degrees" or programs which follow a clear cut path of courses. This system is an intelligent Web-based application for the handling of general advising cases. Authors used Semantic Web technologies for designing this system. Also, in the work (Wen and McGreal, 2007) system allows learners to add preferences of the specialization to their profile and then recommend courses based on their preferences. The proposed system uses a multiple intelligent agents approach and ontology-driven methodology to tackle a dynamic and complex individualized study planning problem. In other research, (Nurjanah, 2016) authors proposed recommender system in adaptive learning which recommends learning materials for learners. This system combines content-based filtering and collaborative filtering approach which is based on the similarity between learners and learners' competencies. The result of the proposed new technique of recommendation showed that it performs well. In paper (Myneni et al., 2013) authors presented an interactive and intelligent learning system for physics education. This environment helps learners master physics concepts in the context of pulleys. System guides learners in problem-solving. Authors used the Bayesian network for pulley setup. In the work (Dolenc and Aberšek, 2015) authors demonstrate the design and evaluation of intelligent tutoring system based on cognitive characteristics of the learner. This model is based on a system for collecting metadata and variables that are vital for the teaching process. The important thing in this research is that system can be adapted to the different level of learner's knowledge and understanding subject matters. Authors in paper (Canales et al., 2007) suggest an adaptive and intelligent Web-based education systems that take into account the individual learning requirements. They proposed three modules for that: an Authoring tool, a Semantic Web-based Evaluation, and Cognitive Maps-based Learner Model. Proposed approach is focused on reusability, accessibility, durability, and interoperability of the learning content.

There are many intelligent tools developed for learning systems which are effective and shows good results. However, there are not much tools or systems which could help learner during his whole lifecycle.

3 ARGUMENTATION FOR THE ILSS

This section proposes the ecosystem of the Intelligent Learning Support System (as presented in figure 1) and describes the main mechanisms that allow learners to get learning content recommendations. The model of the system based on semantic technology. At first, we discuss using Semantic Web in building ILSS. After, the main blocks of the system are presented in sub-chapters. The proposed system is now in the development stage and includes the necessary functionality description required.

3.1 Using Semantic Web in building ILSS

Developing adaptive, intelligent, and personalized systems require comprehensive knowledge in the domain and flexible mechanism for further extension of it as well as the automated mechanism for contextual information retrieval. It will require representation of knowledge, which is widely distributed among different organizations, freely available, and presented in machine readable form. The system should adopt benefits of Semantic Web and linked data technologies, as well as, be facilitated by text mining and natural language processing techniques.

There has been many studies proposed for the intelligent study advising utilizing Semantic Web technologies (Henderson and Goodridge, 2015; Gavriushenko et al., 2015; Ranganathan and Brown, 2007; Nguyen et al., 2008; Jovanović et al., 2007). Authors in (Dicheva et al., 2009; Jovanovic et al., 2007) used ontologies and Semantic Web in curriculum development and in institutional support and adaptation. Semantic Web helps formalize learning object content, shows how to use semantic annotation to interrelate diverse learning artifacts (Jovanović et al., 2009).

Semantic Web aims at integrating various types of information into a single structure, where each semantic data element will meet special syntactic block. It provides a common framework that allows data to be shared and reused across application,

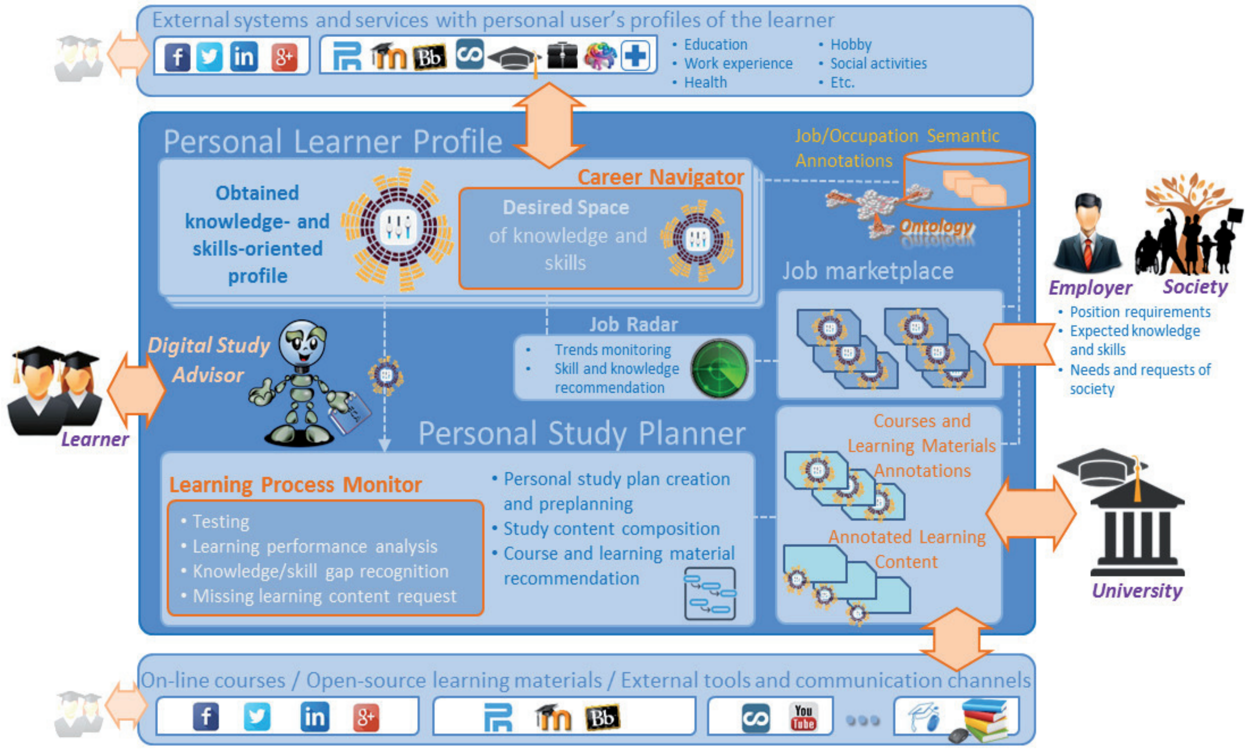


Figure 1: General ecosystem of an Intelligent Learning Support System

enterprise, and community boundaries (Berners-Lee et al., 2001). Semantic Web achieving interoperability among various educational systems and unified authoring support for the creation (Aroyo and Dicheva, 2004). The other branch of the Semantic Web associated with a direction close to the field of artificial intelligence, and it is called the ontological approach.

The ontological approach means a formal representations of a set of concepts within a domain and the relationships between those concepts (knowledge base). Under the knowledge base, we understand complex data structure that stores the knowledge domain. Typically, the knowledge base is represented as a graph in which the vertexes are one unit of learning, that is the simple essence of knowledge (training course, the term, the definition of the formula – depends on an agreement between system's makers that they will assume a basic understanding of the essence). The arcs are seeing as all kinds of links between training units.

Knowledge bases are used for solving problems like creation, transformation, and usage of content while learning. The engineering part includes the creation of the educational plans and their e-learning resources. Management part problems are solved directly during the learning process and are focused on operational adjustment training course depending on the results of the test control of learners' intermediate mastering of the material. Creation of the ontologies as knowledge bases are very useful for tests creation, automation of the evaluation, and management of the training paths based on the results of testing.

To be processed by intelligence (applications, services) and be consumed by end-user in a best optimal way, learning content should be adapted to the digital world of intelligent entities and must be supplied with correspondent semantic annotation. Semantic annotation provides a set of worldwide standards, which helps in operating with heterogeneous resources using common syntax and methods (Uren et al., 2006). Semantic annotation by the rules identifies concepts and their relationships, and it is meant for use by machines (Uren et al., 2006). Semantic annotation was presented and successfully applied in papers (Jovanović et al., 2009; Jovanović et al., 2007), where authors presented ontology-based approach to automate annotation of learning objects and tested it in integrated learning environment for domain of Intelligent Information Systems.

Providing the personalized point of view on content annotation, learner extends a pool of context dependent semantic annotations for the same content. Later, annotations might be transformed into a set of different annotations for the same content on the different clusters of the learners as well as to the wide variety of other contexts. Such clustered semantic annotations will enhance matchmaking for new content retrieving and personalized content recommendation processes.

3.2 Courses and Learning Material Annotation module and Annotated Learning Content module

Usually, educational material and content are presented in the same way for everybody, without taking into account learner's existing knowledge, goals for browsing, preferences, and experience. This issue needs more attention, especially when it comes to e-learning because the learner is supposed to take responsibility for one's studies. The main factors that should be taken into account are learner's knowledge background, age, experience, motivations, professions and goals (Huang et al., 2007).

The common feature of learning support systems is that they have some grade range for learners of different learning levels. Usually, there is some pre-test which can determine learner knowledge and create simple learning path according to the particular level of the learner. A simple adaptation algorithm adds that if the learner makes lots of mistakes, the system returns the learner to the previous level. Otherwise, the system allows the learner to proceed. The threshold of right answers can be eighty percent. Unfortunately, twenty percent of wrong answers can mean that some critical knowledge may be missed, which can interrupt the process in future. It is important to keep track what learner does not exactly know and how this particular knowledge could be improved. The main idea is to find learning gaps which influence the learning process and recommend material to help the learner to fulfill those gaps. Same kind of learning gaps can also occur when the learner is searching for a specific skill and has to take some additional courses that are not included in the curriculum. The targeted courses might have prerequisites, which the person can not meet at the moment.

From the employer's point of view, it is not always possible to recognize the potential of a job seeker from a list of completed courses or degrees. The names of different courses do not always say much about actual content, or even opposite may bring wrong expectations about skillfulness of a person who passed them. Depending on the educational organization, there might be a huge difference between the content of courses with similar names.

It might be more reasonable to see an approximate level of particular skills of a person, to see an expectation of the potential to resolve the certain class of problems. Therefore, all the course descriptions or meta-information should include some explicit reference to skills and knowledge that it is aimed to lead or influence. Thus, having such skills and knowledge oriented annotation of courses and learning materials, it is possible to facilitate a better guidance for a learner.

If we make a straightforward assumption, that the aim of education is to develop certain knowledge and skills to be applied in different occupations, learner chooses education steps (university, program, course, lecture) for a certain goal (profession, career, occupation). Unfortunately, the learner may not be aware enough of the required skills or knowledge for this goal. Therefore, occupations and job descriptions should include skills- and knowledge-oriented annotation as well.

In this work we use semantic annotations with the use of ontologies (as presented in figure 2). Ontologies were used for storing and modeling information about learners, teachers, courses, specializations, universities, careers, job offers, among other things. With ontologies it is possible in future to apply methods of semantic alignment for matching the ontologies (Heath and Bizer, 2011; Berners-Lee, 2006; Henderson and Goodridge, 2015), if there will be some other existing ontologies which are related to our educational system. Nowadays, many tools exist for the creation and matching ontologies which have proven themselves well (Berners-Lee et al., 2001; Shvaiko and Euzenat, 2013). The implementation process was done by Protégé-tool which is the most widely used and offers a complete development environment.

3.3 Personal Learner Profile module

The educational system produces a wide amount of data related to learning, performance, difficulties, and experiences. This information can be stored in a Personal Learner Profile (PLP), which always support changes and keeps the information up-to-date. Interfaces to profile give access to only authenticated learners and systems. The knowledge base would give anonymous data for pedagogical development purposes. Individual learners would be able to access the information related to themselves. Information stored in the personal profile would consist of: study history, work experience, personality, learning plans, learning difficulties, hobby, and social activities.

The PLP would be connected to the external system and services with learner's profiles that connects different systems, services, and databases, for example, the national social and health systems. The ILSS

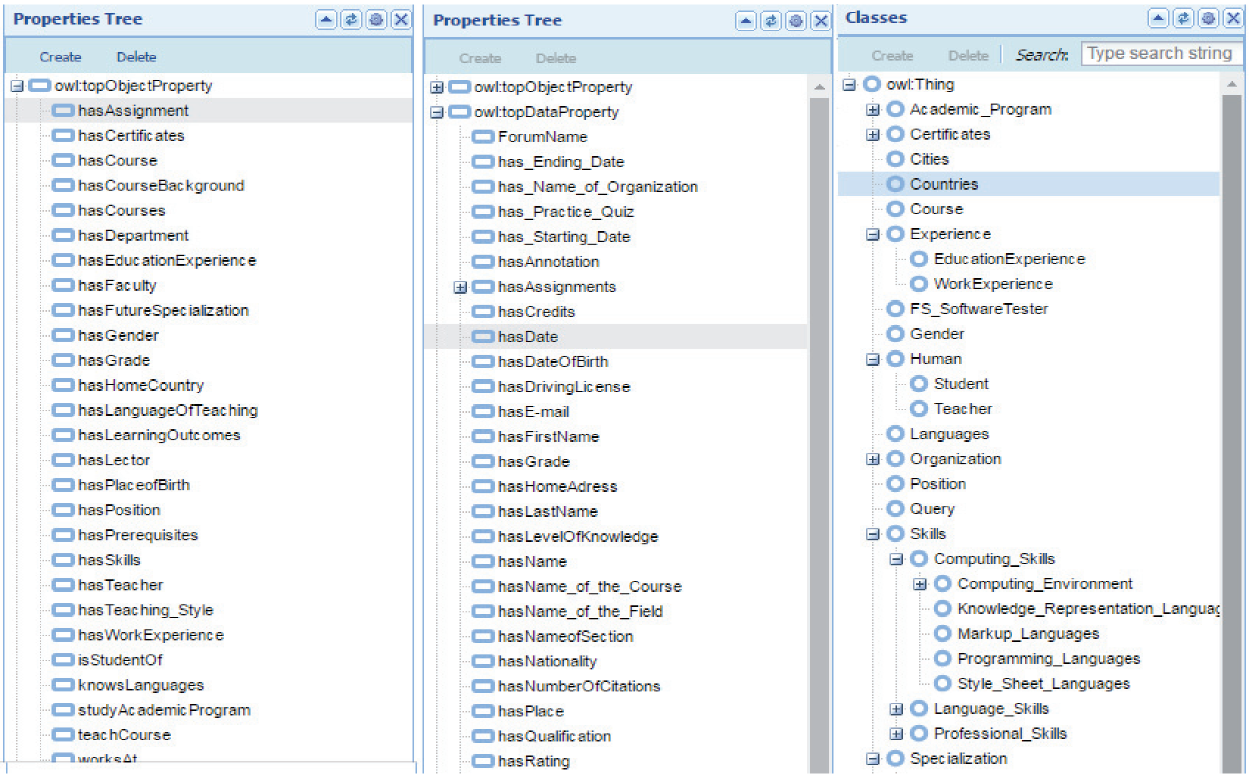


Figure 2: Proposed ontology for an Intelligent Learning Support System

will be developed to analyze the information from systems connected with the digital service channel. The system is based on current research knowledge, and it will have various kind of pedagogical information related to learning theories, practices, environments, and special needs.

The PLP contains current skills- and knowledge-oriented profile of a learner in the form of a vector in multi-dimensional space of knowledge and skills. The *Career Navigator* module helps the learner to build a desired/aimed space of knowledge and skills, based on selected occupations/jobs the learner aims. The Career Navigator is connected to the *Job market place* that contains: job and occupation descriptions published by employers, national needs and requirements, and expectations of society. The Career Navigator is facilitated with proactive "Job Radar" module that monitors/analyses changes in the Job marketplace and informs learner about latest trends. The Career Navigator tool should match curricula, various certificates, job history and other measured skills of a person with the description of profession and job position requirements. Depending on the set goal, the navigator may build a career path to the aimed point, highlight necessary steps and required missing skills for it, offer relevant study programs and online courses to extend personal skills and education. Furthermore, it can show the closest job offers from the labor market, taking into account latest trends in economics development on regional and governmental levels or even worldwide, and offer suitable new profession obtaining programs supported by the government for the temporally unemployed person.

This module should have a possibility to find the program that is more close to the learner's goal of finding the desired job. Learner profile, career, and study content should be described with the same annotation type.

3.4 Personal Study Planner module

The PLP module works as an input for a Digital Study Advisor (DSA), which uses a Personal Study Planner (PSP) to build corresponding study plan and compose the most suitable study content for a learner, depending on the available learning materials. *Learning Materials Space Manager* is connected to various online learning platforms and courses, has access to an open source learning materials, and allows the learner to extend the space with courses provided by any university or education organization learner is enrolled. The DSA analyses the gaps (missing parts) between the current skills-/knowledge-oriented profile of a learner and his/her desired knowledge/skills space and performs a semantic search among annotated courses and learning materials available in Learning Materials Space. Learning Materials Space will also be connected to external tools and communication channels, online courses, open-source learning materials. The PSA also uses the PSP to re-plan initially created study plan on the performance of a learner in the process of his/her studies. With the help of *Learning Process Monitor*, the PSA analyzes the performance of a learner based on real time tests and feedback, recognizes corresponding gaps in knowledge and requests for missing learning materials from the Learning Materials Space Manager.

For the ILSS there will be good to classify the each learning modules. The planning of the learning path will be on the content level, as well as on the content parts level. For the fulfilling knowledge gaps, we have to annotate the whole learning content (courses, study modules, assignments, etc.). Then the system will adapt the content according to learner's feedback.

The ILSS should take into account extracted not only knowledge background of the learner, but also his personal ability for learning new material. According to that, we have to minimize the time for studying and maximize the quality of the learning content. The system should improve the logic of the learning concepts' presentation, taking into account personal specifics of the learner via manipulation with the complexity level of the system. Also, system should analyze learning progress and see what is missing and try to fill missing knowledge by searching the connection between learning modules in the course. This connection search is needed because it will be easy to see on what step of the learning process learner starts to make mistakes and what parts of modules influenced that.

4 DISCUSSION

The best way to produce experts is to accentuate the learner's best abilities and skills, assess the learner's potential and develop it further. That is why we badly need revolutionary methods to facilitate intelligent personalization of study processes and approaches to make innovative education content more attractive and motivational for the learner (Khriyenko and Khriyenko, 2013).

To be able to provide appropriate recommendation or suggestion of relevant content, the system should be aware of learner preferences, experience and knowledge, and goals. Therefore, availability of recently updated learner profile is one of the crucial issues. The process of personal learner profile creation (extension, modification) should not be annoying for the learner, should not be time-consuming and should not demand significant efforts for him/her. The main questions remain the same: how to distinguish relevant parts of a profile to be modified, how to automate and simplify the process of profile management as an integral part of other processes/activities performed by the learner. Therefore, we are going to elaborate content-driven approach for personal learner profile management.

Having semantically annotated content and personalized knowledge/information spaces defined by correspondent semantic profiles of the learners, we may populate the spaces with relevant content. Semantic Web annotations enhance information retrieval and improve interoperability. Unfortunately, human annotators are prone to error, and non-trivial annotations usually require experts in the specific domain. Moreover, that is why annotations should be maintained. If annotations are done cost-effectively, then the future for the technology will not be limited.

The main idea behind proposed system is taking into account knowledge and skills of the person. That is why the whole learning content, courses, and learning materials should be annotated. For the annotation, we chose semantic annotation in the form of ontology. The majority of the universities have to follow the proposed ontology for it is better working and using. In consequence, we can have a full set of courses from different universities, available bachelor's and master's programs with specific requirements (i.e. with certain courses and backgrounds). Learner, using such system will be able to include his full transcript of records, then to find a university, link to courses that are interesting to him, as well as to find the shortest path for getting the higher education by the present knowledge background.

The proposed ILSS makes possible to review and reflect long-term information. One can produce a different kind of analyses related to individual strengths and weaknesses. With the system, one can get additional resources and individual recommendations about learning technologies, materials, courses and support.

A learner can use the system as a virtual career agent. One can get information about available careers and education possibilities, and review personal information against the prerequisites. One can easily produce an application, resume or curriculum vitae and decide what information to include. This will make the process of work applying more straightforward. The objective is that a learner can get complete information related to own educational history and up-to-date information related to work and career possibilities.

For a teacher, the system will provide information about the learner group. The teacher can get up-to-date and individual information about the person's learning, no matter where the learner is coming. This will help the teacher to guide learners and personalize teaching for individual needs. The teacher can produce different kind of analysis about the learners and get information about what methods, materials, and learning environments could be beneficial for them. The objective is that a teacher can get relevant and up-to-date information about learners, current research knowledge from the system, help with difficulties, support practices and methods related to learning.

This plan tries to fulfill the needs of modern digital and global society where learning and teaching are not necessarily bounded to single institutions. The concept of individual know-how will extend from degrees to a complete picture where formal education is only one part. Further study and experiments will consist of an extension of ontology for the University of Jyväskylä. This University has an in-house developed, integrated study information system *Korppi*[1], which already has some annotation about courses, and study modules. Because of that, it would be easy to transfer presented information in the well structured knowledge base. As well as at the University of Jyväskylä is possible to transfer courses and credits from other universities or online learning platforms.

5 CONCLUSIONS

Usually, computer based training or computer aided learning does not use the model of the learner's knowledge and does not take into account personality of the learner, but it just uses traditional instructional methods (Phobun and Vicheanpanya, 2010). The combination of adaptation, intelligence and personalization could improve the learning process.

In Web-based learning environments, it is an advantage to monitor the learner's knowledge level and automatically adjusts the content for each learner to improve the education process and make learning individual for each learner or group of learners. Since the Internet offers a vast amount of information, it is important to help the teacher in the creation of the materials most suitable for education; and as well help in finding the most relevant content and convert it to comprehensive information. Also, a big privilege of these Web-based learning environments is to help learners with selecting courses and study programs basing on their objectives, which will be beneficial in their further career development.

This paper describes the issues of the lifelong learning, and how to make it more personalized, supported by technology. Paper presents Intelligent Learning Support System and describes it is main modules.

An ontology to serve as the knowledge base of ILSS was developed. All semantically annotated information will be stored there for future SWRL-formulated rules which allow extraction of new facts from existing knowledge. The developed ontology is very flexible, and it is possible to modify data easily.

ILSS is connected to the different open-source learning materials, external tools and communication channels, and online courses. Also, external systems and services with personalized user profiles can provide better guidance for a learner in a personalized manner, as well as suggest new relevant content or events which could help to fill learner's knowledge gaps.

The proposed system might be very useful for universities and companies. It could also help in monitoring changes in labor market and modification of study paths according to the called-for competence, as well as to help in career planning.

[1] https://www.jyu.fi/itp/en/korppi-guide

REFERENCES

Aroyo, L. and Dicheva, D. (2004). The new challenges for e-learning: The educational semantic web. *Educational Technology & Society*, 7(4):59–69.

Bendakir, N. and Aïmeur, E. (2006). Using association rules for course recommendation. In *Proceedings of the AAAI Workshop on Educational Data Mining*, volume 3.

Berners-Lee, T. (2006). Linked data. design issues for the world wide web. *World Wide Web Consortium. http://www. w3. org/DesignIssues/LinkedData. html*.

Berners-Lee, T., Hendler, J., Lassila, O., et al. (2001). The semantic web. *Scientific american*, 284(5):28–37.

Brusilovsky, P. and Peylo, C. (2003). Adaptive and intelligent web-based educational systems. *International Journal of Artificial Intelligence in Education (IJAIED)*, 13:159–172.

Canales, A., Peña, A., Peredo, R., Sossa, H., and Gutiérrez, A. (2007). Adaptive and intelligent web based education system: Towards an integral architecture and framework. *Expert Systems with Applications*, 33(4):1076–1089.

Chen, C.-M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 51(2):787–814.

Dicheva, D., Mizoguchi, R., and Greer, J. E. (2009). *Semantic web technologies for e-learning*, volume 4. Ios Press.

Dolenc, K. and Aberšek, B. (2015). Tech8 intelligent and adaptive e-learning system: Integration into technology and science classrooms in lower secondary schools. *Computers & Education*, 82:354–365.

García-Barrios, V. M., Mödritscher, F., and Gütl, C. (2005). Personalisation versus adaptation? a user-centred model approach and its application. In *Proceedings of the International Conference on Knowledge Management (I-KNOW)*, pages 120–127.

Gavriushenko, M., Kankaanranta, M., and Neittaanmäki, P. (2015). Semantically enhanced decision support for learning management systems. In *Semantic Computing (ICSC), 2015 IEEE International Conference on*, pages 298–305. IEEE.

Gavriushenko, M., Khriyenko, O., and Porokuokka, I. (2016). Adaptive vocabulary learning environment for late talkers. In *CSEDU 2016: Proceedings of the 8th International Conference on Computer Supported Education. Vol. 2, ISBN 978-989-758-179-3*. SCITEPRESS.

Heath, T. and Bizer, C. (2011). Linked data: Evolving the web into a global data space. *Synthesis lectures on the semantic web: theory and technology*, 1(1):1–136.

Henderson, L. K. and Goodridge, W. (2015). Adviseme: An intelligent web-based application for academic advising. *International Journal of Advanced Computer Science & Applications*, 1(6):233–243.

Huang, M.-J., Huang, H.-S., and Chen, M.-Y. (2007). Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach. *Expert Systems with Applications*, 33(3):551–564.

Jeong, H.-Y., Choi, C.-R., and Song, Y.-J. (2012). Personalized learning course planner with e-learning dss using user profile. *Expert Systems with Applications*, 39(3):2567–2577.

Jovanovic, J., Gasevic, D., Brooks, C., Devedzic, V., Hatala, M., Eap, T., and Richards, G. (2007). Using semantic web technologies to analyze learning content. *IEEE Internet Computing*, 11(5).

Jovanović, J., Gašević, D., and Devedžić, V. (2009). Tangram for personalized learning using the semantic web technologies. *Journal of emerging technologies in web intelligence*, 1(1):6–21.

Jovanović, J., Gašević, D., Knight, C., and Richards, G. (2007). Ontologies for effective use of context in e-learning settings. *Educational Technology & Society*, 10(3):47–59.

Jyothi, N., Bhan, K., Mothukuri, U., Jain, S., and Jain, D. (2012). A recommender system assisting instructor in building learning path for personalized learning system. In *Technology for Education (T4E), 2012 IEEE Fourth International Conference On*, pages 228–230. IEEE.

Khriyenko, O. and Khriyenko, T. (2013). Innovative education environment and open data initiative: Steps towards user-powered society-oriented systems. *GSTF Journal on Computing (JoC)*, 3(3):31.

Leung, C. M., Tsang, E. Y., Lam, S., and Pang, D. C. (2010). Intelligent counseling system: A 24 x 7 academic advisor. *Educause Quarterly*, 33(4):n4.

Myneni, L. S., Narayanan, N. H., Rebello, S., Rouinfar, A., and Pumtambekar, S. (2013). An interactive and intelligent learning system for physics education. *IEEE Transactions on learning technologies*, 6(3):228–239.

Nguyen, T. B., Nguyen, D. N., Nguyen, H. S., Tran, H., and Hoang, T. A. D. (2008). An integrated approach for an academic advising system in adaptive credit-based learning environment.

Nurjanah, D. (2016). Good and similar learners' recommendation n adaptive learning systems. pages 434–440.

Phobun, P. and Vicheanpanya, J. (2010). Adaptive intelligent tutoring systems for e-learning systems. *Procedia-Social and Behavioral Sciences*, 2(2):4064–4069.

Ranganathan, G. R. and Brown, J. A. (2007). The use of ontologies and rules to assist in academic advising. In *International Workshop on Rules and Rule Markup Languages for the Semantic Web*, pages 207–214. Springer.

Shvaiko, P. and Euzenat, J. (2013). Ontology matching: state of the art and future challenges. *IEEE Transactions on knowledge and data engineering*, 25(1):158–176.

Uren, V., Cimiano, P., Iria, J., Handschuh, S., Vargas-Vera, M., Motta, E., and Ciravegna, F. (2006). Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: science, services and agents on the World Wide Web*, 4(1):14–28.

Wen, F. L. S. L. D. and McGreal, F. Z. K. R. (2007). e-advisor: A multi-agent system for academic advising.

Werghi, N. and Kamoun, F. K. (2009). A decision-tree-based system for student academic advising and planning in information systems programmes. *International Journal of Business Information Systems*, 5(1):1–18.

PV

SUPPORTING INSTITUTIONAL AWARENESS AND ACADEMIC ADVISING USING CLUSTERED STUDY PROFILES

by

Mariia Gavriushenko, Mirka Saarela, Tommi Kärkkäinen 2017

9th International Conference on Computer Supported Education (CSEDU 2017), Porto, Portugal

Supporting Institutional Awareness and Academic Advising Using Clustered Study Profiles

Mariia Gavriushenko, Mirka Saarela and Tommi Kärkkäinen
Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland
magavriu@student.jyu.fi, mirka.saarela@jyu.fi, tommi.karkkainen@jyu.fi

Keywords: Academic Advising, Learning Analytics, Robust Clustering.

Abstract: The purpose of academic advising is to help students with developing educational plans that support their academic career and personal goals, and to provide information and guidance on studies. Planning and management of the students' study path is the main joint activity in advising. Based on a study log of passed courses, we propose to use robust, prototype-based clustering to identify a set of actual study path profiles. Such profiles identify groups of students with similar progress of studies, whose analysis and interpretation can be used for better institutional awareness and to support evidence-based academic advising. A model of automated academic advising system utilizing the possibility to determine the study profiles is proposed.

1 INTRODUCTION

The credit-based system is used to characterize the requirements and progress of a student in many learning environments. Availability of personalized support is very important constituent of a learner's success (Nguyen et al., 2008). Academic advising (AA) is an iterative collaboration process between student, academic adviser, and academic institution, to tackle the student retention. Advisers provide versatile assistance to the students during their studies, making the educational experience relevant and supported.

AA activity has a long history dating back to 1870s (Tuttle, 2000). Advising starts, when a student becomes enrolled in higher education, and finishes when the degree has been completed. The purpose of academic advising is to ensure that the students carry out the required studies to graduate. The central activity for this purpose is to support study planning, especially at the beginning of the academic life. Depending on the organizational culture, especially the stability or dynamicity of the course schedule, an academic adviser either needs to ensure that predefined study plan is being followed or that a student knows all the relevant study possibilities. Email, social media, web and wikipages etc. provide means to share the necessary information with the students. However, typically face-to-face discussions take place either regularly (e.g., at the beginning of a semester or academic year) or by students' or advisers' request. For more personalized support, an adviser should know when a student is in need of a study advice discussion and what is the precise status of the studies.

Preliminary recommendations to do certain, especially compulsory, courses to proceed normally with the major subject studies are provided by departments and advisers. More precisely, for example at the University of Jyväskylä (JYU), all students are required to prepare an electronic personal study plan with the academic adviser from the home department. Advising is organized according to the *satellite model* referring to the distributed responsibility of academic units (Tuttle, 2000). The *engagement model* between advisee and adviser characterize the principal spirit of counselling (Feghali et al., 2011).

In JYU, the study plan and the completed studies create the starting point for a study plan assessment discussion. However, especially in computer science, the actual number of studies that have been made during an academic year are typically less than recommended (Saarela and Kärkkäinen, 2015a). Hence, it is very common that the actual study path deviates from the recommendations and plans, and in such a case an advising intervention is needed. But how does an individual student and especially an individual study adviser know, what is the relation between students' realized study path and that of the peer students? Thus, could we, instead of comparing against the predefined plans, advise students based on evidence from the actual study paths of other similar students?

Therefore, the purpose of this article is to propose using a learning analytics method (Chatti et al., 2012), more precisely robust clustering (Äyrämö, 2006; Saarela and Kärkkäinen, 2015a), to create groups of actual profiles of students concerning their studies. Such profiles summarize the different typical accumulations of completed studies, increasing the general awareness of the common study flows. One can explicitly link an individual student to student peers with similar study path profile in the same institutional environment. This allows an adviser to plan a possible intervention adaptively for a larger pool of students instead of following each one individually.

Creating general study profiles of students can help departments in their assessment and planning of when (and how) they provide the courses, especially the compulsory ones (Saarela and Kärkkäinen, 2015a). By automating general student profiling it is possible to provide essential support for adaptive on-line advising along the lines suggested, e.g., in (Nguyen et al., 2008; Henderson and Goodridge, 2015). Individual student's perspective, from self-regulated learning and study planning point of view without academic advising interface, was thoroughly addressed in (Auvinen, 2015) (see also (Auvinen et al., 2014)).

The contents of this article is as follows: after the introduction, we provide background on academic advising and personalized student support in Section 2. Then, in Section 3, we describe the data and the robust clustering method, and introduce three cases to construct student profiles to support academic advising. In Section 4 features and a model of an Academic Adviser system with automated mechanisms in it are proposed. The work is concluded in Section 5.

2 BACKGROUND

Academic advising is a collaborative process in which adviser and advisee enter a dynamic relationship where adviser helps advisee to enhance the learning experience by helping in making academic decisions (Henderson and Goodridge, 2015). The decision support could be made by analysis of student's records, as well as some external factors like interests, goals, academic capabilities, schedules etc. (Noaman and Ahmed, 2015). Developmental Advising means helping students to define and explore academic and career goals and pathways, as well as to develop problem-solving and decision-making skills; Prescriptive Advising, which is the more traditional advising model, is mainly concentrated on providing the information to the students according to their academic program, progress, academic policies, course selection, etc.; Intrusive Advising refers to contacting with student in critical periods like first year of study before the declaring major, graduation period, or when students are at-risk or they are high-achieving students (Noaman and Ahmed, 2015).

Next we briefly summarize a pool of directly AA related work that was identified through a non-systematic search. Our main concern here is to illustrate the strong link between the needs and practices of AA and the general utility of student profiling.

2.1 On Academic Advising

Student voices on AA were raised in only some articles. El-Ansiri et al. (Al-Ansari et al., 2015) used questionnaire to study student satisfaction and support-seeking patterns among dental students in Saudi-Arabia. Very low (only 7.6%) primary utility help rate of advisers in the academic matters was encountered. Even if the advisers were available when needed, they were not able to provide the most relevant information, e.g., on important dates and courses. Hence, up-to-date course and timetable information seems to be a prerequisite for AA, which is handled by the in-house developed, integrated study information system *Korppi*[1].

The pedagogical side of AA was also focused rarely. In the work (Drozd, 2010) author studied academic advisers through the lens of transformational leadership, i.e. how advisers can create a connection to students that positively influence their study paths (by increasing and inspiring study motivation and engagement/commitment in studies through individual and intellectual consideration). A questionnaire for undergraduate students strengthened the importance of transformational leadership activities in adviser-student communication and collaboration, independently from the student's characteristics. The lack of time for individual counselling efforts that was visible in most of the reviewed articles here was not emphasized in (Drozd, 2010). Dougherty (Dougherty, 2007) studied academic advisers from those students' perspective who are doing very well in their studies. These students are called high-achieving students. Authors address the need for the investigation of unique characteristics of these students.

Technical support for AA has been considered in many articles. The availability of extensive information on courses to support automatization of AA was emphasized in (Biletskiy et al., 2009). The authors proposed course outline data extractor application, which helps in recognizing similar or comparable courses between different institutions, also help-

[1] `https://www.jyu.fi/itp/en/korppi-guide`

ing both students and academic advisers to keep track of the variety of topics that i) have been covered in the completed studies, ii) should be covered to complete minor or major subject modules or the actual degree. The authors in (Nguyen et al., 2008) proposed an integrated knowledge-based framework based on semantic technology that supports computer-based (automatic) e-Advising on the suitable courses for the students. Naturally individual learning history data provide the starting point for the system and, for this purpose, the authors implemented and tested a data integration tool.

The high workload of academic advisers, especially due to individual but many times recurrent handling of basic issues with multiple students in a hurry, was addressed in (Henderson and Goodridge, 2015), with the proposition of an intelligent, semantic, web-based application to assist decision making and automatization of repetitive counselling tasks. Core of the system consisted of rule-based inference engine, which mapped student profile with the study program profile and organizational rules, to provide automatic suggestions on the courses to be enrolled in the upcoming semester. In the preliminary evaluation, a positive feedback of the system was obtained, although the main limitation of suitability to only study programs which follow a clear, predefined study path of courses, was recognized. With very similar aims and functionality, another web-based on-line adviser was described in (Feghali et al., 2011). This system was also evaluated positively when compared to the current advising system. The authors emphasized that such a tool only supports and does not replace a human academic adviser.

Conversational, fully autonomous agent supporting AA dialogs using natural language processing (NLP) were suggested in (Latorre-Navarro and Harris, 2015). The proposed system contained an extensible knowledge base of information and rules on academic programs and policies, course schedules, and a general FAQ. NLP performance of the proposed system was evaluated positively. Also, the similar multi-agent approach was suggested in (Wen and McGreal, 2015) for AA. This approach helps tackling a dynamic and complex individualized study planning and scheduling problem. As well as in (Al-Sarem, 2015) was proposed a decision tree model for AA affairs based on the algorithm C4.5. The output is evaluated based on Kappa measure and ROC area. The main conclusion was made that the difference between the registered and gained credit hours by a student was the main attribute that academic advisers can rely on (Al-Ansari et al., 2015).

2.2 On Personalization of Student Support

In general, many researchers have paid attention to developing *e-learning systems with personalization*, and the most common aspect in these system is the *creation of the personalized learning path* for each individual student or group of students. Most of personalized systems consider learner preferences, interests and browsing behaviors, because it will help to provide personalized curriculum sequencing service (Huang et al., 2007). In the study (Chen et al., 2005) authors proposed a personalized e-learning system which is based on Item Response Theory (PEL-IRT). This system is considering course material difficulty and learner ability, to provide individual learning path for learners. Learner's ability estimation was based on an explicit learner's feedback (the answers of learners to the assigned questionnaires). The system appeared mostly like a recommendation system of the courses for the learners. Authors in (Huang et al., 2007) proposed a genetic-based curriculum sequencing approach and used case-based reasoning to develop a summative assessment. The empirical part indicated that the proposed approach can generate appropriate course materials for learners taking into account their individual requirements. Later, in (Chen, 2009a), the authors developed a personalized web-based learning system grounded on curriculum sequencing based on a generated ontology-based concept map, which was constructed by the pre-test result of the learners. Optimization problem for modeling criteria and objectives for automatic determination of personalized context-aware ubiquitous learning path was suggested in (Hwang et al., 2010). This learning model not only supports learners with alternative ways to solve problems in real-world situations, but also proposes more active interaction with the learners. Authors in (Werghi and Kamoun, 2009) proposed Decision Support System for student advising based on decision tree for an automated program planning and scheduling. The proposed approach takes into account prerequisite rules, the minimum time (minimum number of terms), and the academic recommendations. The adaptive course sequencing for personalization of learning objectives was suggested in (Idris et al., 2009) using neural networks, self organizing maps and the back-propagation algorithm.

A very closely related work to ours was reported in (Sandvig and Burke, 2005). Authors proposed a case-based reasoning paradigm which is based on the assumption that similar students will have similar course histories. The system used the experience and history of graduated students in order to propose potential courses for the students. Unfortunately, this approach required matching between students' histories. Also, similar case-based reasoning was used by (Mostafa et al., 2014) for developing a recommendation system for a suitable major to students based on comparison of the student information and similar historical cases.

3 CREATION OF STUDY PATH PROFILES USING ROBUST CLUSTERING

3.1 Data

To illustrate the proposed approach, we utilized real study records of the Bachelor (BSc) and Master (MSc) students majoring in Mathematical Information Technology (which is comparable to a major in Computer Science at other universities) at the University of Jyväskylä (JYU/MIT). IT administration at the University has recently created a data warehouse of passed courses by all the students, which can be utilized by the departments. On the other hand, the electronic study plan system does not provide direct interface for larger student groups, so both from accessibility and evidence-basedness points of view, we focus on analyzing the real study log of the passed courses. The log was anonymized, keeping student IDs as keys, covering the four calendar years 2012–2015. Note that students can start their studies in the beginning of September (autumn term) or January (spring term). Hence, the original study registry log included a heterogeneous set of BSc and MSc students who had started their studies either before 2012 or in the beginning of spring or autumn terms during 2012 – 2015.

The whole study log contained 15370 passed courses by 1163 different students on 1176 different course IDs. There were 942 male students (81%), with mean amount of studies made 59.9 ECTS. Only 221 female students (19%) were identified, with mean amount of studies made 57.0 ECTS. Hence, most of the students in the log were either in the beginning of their studies or progressing very passively and slowly.

Figure 1 shows the discrete density distribution of the size of the passed courses. According to the figure, 5 ECTS and 3 ECTS are the two most common sizes of the courses, the former covering around 30% of the studies. Moreover, there are a lot of small courses (1 – 6 ECTS) with the exception of the MSc thesis, 30 ECTS. Teaching in JYU is organized for four periods during one academic year (plus the summer semester) in such a way that a course of ca. 5 ECTS can fit to one period. We conclude that because the passed courses represent both major and minor subject studies, division of the overall learning objects as courses is not optimal. This observation is the first example on how summarization of study log data provides visibility and feedback to the organization. During the course of writing this article, we also found out that the instructions of JYU for preparing the next curriculum for 2017–2019 include strong recommendation to decrease the number of courses with only a few credits.

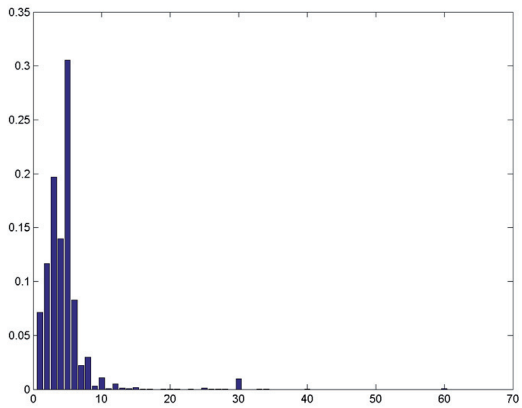


Figure 1: Probability of size of a course.

Next we aggregated how many credits per semester each student had made. Similarly to (Saarela and Kärkkäinen, 2015a), each calendar year was divided into two semesters: the spring term (from January to June) and the autumn term (from July to December). However, since usually only a few courses are completed during the summer (this is illustrated in (Saarela and Kärkkäinen, 2015a)), it was reasonable to divide the calendar year into only two parts for further analysis.

In what follows, we profile, analyse and compare two students cohorts: those who started their studies in the beginning of the autumn term 2012 (A2012) or 2013 (A2013). Hence, for A2012 we end up with 8 and for A2013 with 6 integer variables representing the aggregated amount of credits on half-a-year scale. Since the students have progressed in their studies very differently and many of them have not been active during all the semesters of interest, both of the data sets are very sparse containing a lot of missing values (Saarela and Kärkkäinen, 2015a). This is the key property that is taken into account in the profiling approach that is described next.

3.2 Robust Clustering Method

As already explained, our goal is to assist the academic advisers by recommending suitable courses for students based on passed courses of (possibly more advanced) students with similar study path. For this, we need to identify general profiles of similar students and this is, precisely, the purpose of clustering. Partitional (or representative-based (Zaki and Meira Jr, 2014)) clustering seems to be the right family of clustering methods to choose from because it assigns each observation to exactly one cluster, which is represented by its most characteristic point, the cluster centroid, which represents the common profile. Within a cluster, distances of observations to the prototype determine the most typical or representative members of a cluster. Thus, instead of following many different student profiles, the academic adviser can just follow the most common profiles to get an overview of the whole cohort.

Generally, partitional based clustering algorithms consist of an initialization step, in which the initial centroids of each cluster are generated, and iterations of two steps where (i) each observation is assigned to its closest centroid, and (ii) the centroid of each cluster is recomputed by utilizing all observations assigned to it. The algorithm stops when the centroids remain the same over two iterations. The most popular and most applied partitional clustering algorithm is the *k-means* (Jain, 2010), also in learning analytics studies (Saarela and Kärkkäinen, 2017). This algorithm works very well for full and approximately normally distributed data since the sample mean is the most efficient estimator for samples that are drawn from the normal distribution. However, the sample mean is highly sensible to all kinds of outliers (Huber, 2011) as well as missing values, which can be characterized as special types of outliers. Also for a non-symmetric (skewed) distribution, the sample mean is not necessarily the most efficient estimator and other location estimates might be preferable (Sprent and Smeeton, 2016). Moreover, as explained in (Saarela and Kärkkäinen, 2015a), the quantization error for the integer-type variables like here has uniform not gaussian distribution.

The spatial/geometric median is a robust nonparametric location estimate, which remains reliable even if half of the data is contaminated (Sprent and Smeeton, 2016). Mathematically, the spatial median is the Weber point that minimizes the (nonsquared) sum of the Euclidean distances to a group of given points $\{\mathbf{x}_i\}, i = 1, 2, \ldots n$:

$$\arg\min_{\mathbf{c}} \sum_{i=1}^{n} \|\mathbf{x}_i - \mathbf{c}\|.$$

Although the basic concept is easily understood and has been extensively discussed in the literature (albeit under various names, see (Drezner and Hamacher, 2001)), its computation is known to be difficult.

In (Äyrämö, 2006), the difficulty of computing the spatial median during partitional clustering was solved with the SOR (Sequential Overrelaxation) algorithm (see (Äyrämö, 2006) for details). Moreover, in the implementation of the resulting *k-spatial-medians* clustering algorithm, only the available (i.e. not-missing) data is taken into account when the centroid is recomputed.

To sum up, all of these above discussed properties – most importantly, the robustness to missing data and the fact that every cluster is represented by a centroid – make the *k-spatial-medians* clustering very suitable for creating student's general study profiles. The fact that such a clustering approach works very well for sparse educational data has been previously shown in (Saarela and Kärkkäinen, 2014; Saarela and Kärkkäinen, 2015b; Saarela and Kärkkäinen, 2015). The initialization of the robust clustering method was realized similarly as in (Saarela and Kärkkäinen, 2015): We started with multiple repetitions of *k-means* for the complete data – without missing values – and then, applied *k-spatial-medians* to the best of those results.

3.3 Clustered Student Profiles

Similarly to the earlier work in (Saarela and Kärkkäinen, 2014; Saarela and Kärkkäinen, 2015b; Saarela and Kärkkäinen, 2015a; Wallden, 2016), we apply four different internal cluster validation indices to determine the number of clusters: Knee Point (KP) of the clustering error, Ray-Turi (RT), Davies-Bouldin (DB), and Davies-Bouldin$*$ (DB$*$). All the computations here were carried out in the Matlab-environment, using own implementations of all the algorithms.

From the two student groups A2012 and A2013, we include in clustering only still active students, i.e. those who have made credits during the autumn term 2015 (the last one analyzed). Furthermore, we restrict ourselves to those students for whom over half of the variables are available (Sprent and Smeeton, 2016). This means that the 47 analyzed students in A2012 have made studies during at least four out of the seven possible semesters (including the last one) and the 76 students in A2013 at least in three out of the five semesters. Because of the anonymity, we obtained further assistance in relation to the metadata and interpretation of the clusters from the Study Amanuensis of the Department (Study Amanuensis, 2016).

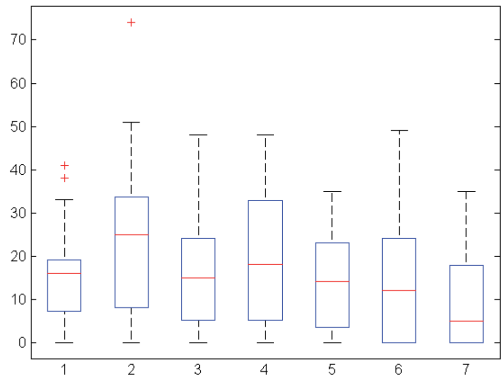


Figure 2: Boxplot for A2012.

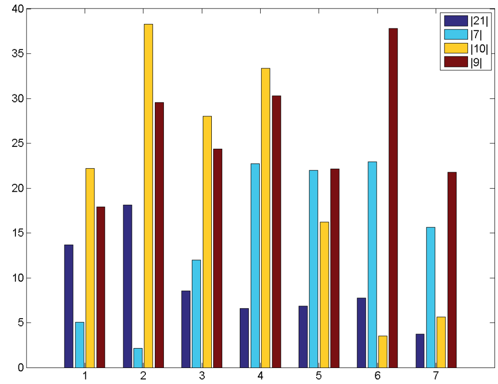


Figure 3: Credit accumulation prototypes for A2012.

A2012

The boxplot in Figure 2 shows the large variability in the study accumulations both within semesters and between semesters. We see the larger accumulations in the spring terms during the first two years, and a slightly decreasing overall trend after that. There are always exceptional students who have made much more studies than their peers.

KP, DP, and DP∗ indicated four clusters and RT had also local minimum there, so we choose to analyze four different general study progress profiles. The profiles for A2012 are depicted in Figure 3, where the size of the cluster is given in the top-right corner. The profiles are sorted in the ascending order with respect to the total number of credits.

The main group of 21 students in the first cluster illustrate a potential start of the studies in the first year, with strong passivation after that. They have obtained prototypically 65 ECTS until the end of 2015. Based on (Study Amanuensis, 2016), by a closer look on the 8 students from the cluster closest to the centroid, these are all older BSc and MSc male students (born before 1990). They are either distant students studing while working or have completely chosen to change their orientation from an earlier occupation and already finished degree. The difficulties in studies and reasons of such a behavior, for a similar adult student profile, were thoroughly discussed in the earlier work (Kaihlavirta et al., 2015) from the same context (department) than here.

The second group of only 7 students, who generally obtained 103 ECTS, shows opposite behavior: very slow start in the first year, activating to an appropriate level then. Three most characteristic students here were young males, who were involved in the military service during the first study year. This complete explains the observed behavior.

The third group of 10 students, who generally obtained 147 ECTS, did their studies very actively for the first 4–5 semesters. Analysis of the three most characteristics students revealed two young and one older male students who either took job or became active in student organizations during the third year of the studies.

The fourth profile with 9 students, altogether 184 ECTS in general, illustrates that a good start on the study activity carries over the semesters. Three mostly characteristics students were again all males, one MSc student and two BSc students. Note that similar finding on the importance of active start in an individual course level was given in (Saarela and Kärkkäinen, 2015a).

Students who are mostly in need of academic advising are the ones in the first cluster. They can be identified either in the beginning of their studies or after the second semester, because even if still making studies, their accumulation is much less than in the third and fourth cluster. Their characterization also suggests the department to rethink the study entrance criteria.

A2013

For A2013 all cluster indices suggested three profiles, which are illustrated in Figure 5. This and the fact that there are now one profile less than in A2012 suggests more stable organization of the curriculum. Also the boxplot in Figure 4 supports such finding, especially showing smaller variability in the obtained credits between the autumn and spring terms compared to A2012 in Figure 2.

Student group in the need of intrusive academic advising consists of those 23 students with smallest accumulation of credits. These students start and continue very slowly in their studies, although the level of activity was increasing in the fourth and fifth

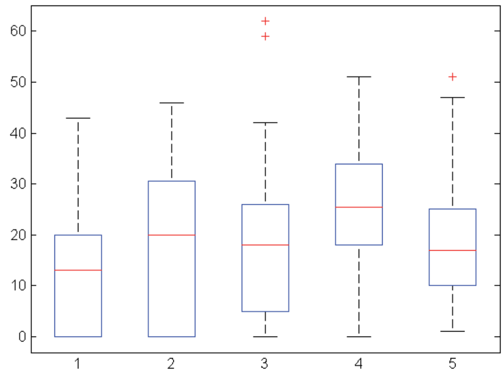


Figure 4: Boxplot for A2013.

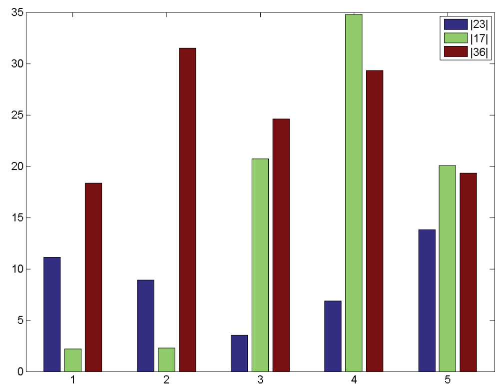


Figure 5: Credit accumulation prototypes for A2013.

semesters. Their general ECTS accumulation after five semesters was 44 ECTS. Analysis of the five most representative students revealed two older male students (birth year before 1990), two males with indications of military service, and a female student. According to (Study Amanuensis, 2016), especially the younger students showed signs of low self-regulation during academic advising sessions.

The second profile of 17 students, completing typically 80 ECTS, showed similar behavior to the second profile in A2012: the minimal first year is raised to a good level of study activity later. A closer look on the five most representative students showed young, two female and three male students. Four of these had identified themselves as a non-active student during the first study year, again mostly due to the military service of the young male students.

The third profile of 36 most active students, accomplishing 123 ECTS typically, showed similar overall behavior than the fourth profile in A2012. The first semester is slightly smaller but then the study path proceeds in the desired way. Recapitulation of the meta data of five most representative students showed five male students, of whom three were oriented towards game programming and development - the most recent study line of the department.

We note that even if the boxplot in Figure 4 indicated more stable study path with respect to autumn and spring semesters, the two profiles of truly active students still illustrated larger study accumulations in the spring than in the autumn. These findings are, however, mostly explained by the longer calender time for the two periods in the spring term compared to the autumn term – a general peculiarity of the Finnish higher education system.

The similarities and differences between the two sets of profiles just discussed emphasize the importance of the use of evidence-based information in academic advising. On one hand, there are repetitive profiles of students proceeding in their studies well or slowly. The latter ones needs to be detected and supported in an intrusive manner in academic advising. The home department responsible for major subject studies and the other departments providing minor subject studies should be informed about the found hindrances of the study paths. In the case analyzed here, there is a clear change of study accumulation profiles from A2012 to A2013, which suggests that the organization of courses, the capabilities of students, and/or their support through academic advising have improved in the educational organization under study.

4 PROPOSITION OF A SYSTEM MODEL

As shown, it is important to follow the actual progress of the students in their studies. There might be no need for an advising intervention, but if so, one should automatically notify the students and the study counselling on the deviations in the study path. The problems of not passing courses and not following suggested study plans usually also call for organizational considerations whether learner ability and the difficulty level of the recommended curriculum are matched to each other properly (Huang et al., 2007).

L. K. Henderson and W. Goodridge (Henderson and Goodridge, 2015) proposed the following division of the *computer supported AA systems* (i.e., Academic Advisers): Basic Computerized Systems, Advanced Automated Systems, and Intelligent Advanced Automated Systems. In this paper, we mostly focus on Advanced Automated and Intelligent Advanced Automated Systems, because these kind of systems can significantly reduce the human adviser's work and the student load in supporting the study process.

As can be concluded from Section 2.1, earlier studies have mostly concentrated on research prototypes which focus only on few main components or tool support for existing learning management systems. Taking into account that user modeling is one of the key factors for including personalization into the learning system, many researchers used ontologies for learners' models, because ontologies have many advantages for creation of user models (Idris et al., 2009; Chen, 2009a; Leung et al., 2010; Nguyen et al., 2008; Biletskiy et al., 2009; Henderson and Goodridge, 2015).

Data-mining techniques have also been applied to the learning environments in order to track users' activities, extract their behavior profiles and patterns, and analyze the data for future improvement of the learning results, as well as for identifying types of learners (Minaei-Bidgoli, 2004). Mostly, for developing personalized learning plan, researchers used decision tree search, heuristic algorithms, genetic algorithms, item response theory and association rules. Also, many studies used semantic web technologies, neural networks and multi-agent approach. Most of the previous studies on personalized learning path generation schemes have mainly focused on guiding the students to learn in the digital world; i.e. each learning path represents a set of digitalized learning objects that are linked together based on some rules or constraints (Liu et al., 2008). While determining such digitalized learning paths, the learning achievements, on-line behaviors or personalized features (such as learning style) of individual students are usually taken into consideration (Schiaffino et al., 2008; Chen, 2008; Chen et al., 2008; Chen, 2009b; Chen et al., 2005).

As reviewed in Section 2.1, many suggestions for intelligent software and information system support of AA have been given. Many studies describe the creation of intelligent learning systems that can make a curriculum sequencing more flexible for providing students with personalized and adaptive study support services (Fung and Yeung, 2000; Lee, 2001; Brusilovsky, 1998; Lee, 2001; Papanikolaou et al., 2002; Tang and McCalla, 2005). Universities are more and more looking into developing self-service systems with intelligent agents as an addition or replacement for the labor-intensive services like academic advising. For example, The Open University of Hong Kong has developed an intelligent on-line system that instantly responds to enquiries about career development, learning modes, program/course choices, study plans, and graduation checks (Leung et al., 2010).

However, the institutional starting point concerning available digital information, especially for the web-based systems that have been proposed, seems to vary a lot. Some systems start and focus on providing easy access to course and degree requirements information whose availability is to be assured first. On the other hand, we might start from the situation where we can readily access most of the relevant data: i) course information with basic contents, learning goals, assessment methods, acceptance criteria, schedule and location, teachers and lecturers etc.; ii) individual, anonymous study records on passed courses and completed studies. (Note that reliable information on student admission is currently not directly available in the organization under consideration).

4.1 Proposed System Architecture

This subsection describes the novel system architecture for AA and automatic feedback based on recognized student group profiles which are obtained by using clustering. We also present an overview of the AA process as both manual and automated process.

The architecture of the proposed system's model for the AA is presented in Figure 6. The system has two main databases: learner profile database and curriculum database. Learner profile database stores learner's data about studies, assessment results, timetables of completed studies, etc. Curriculum database stores information about compulsory courses, other courses, timetables, etc. The academic advising system's part consists of several blocks like linking individual students to their peers with similar study path profile together with the recommendation block and planning block.

Based on the system architecture, the details of system's main functionality read as follows:

1. Collection of learner's personal information.

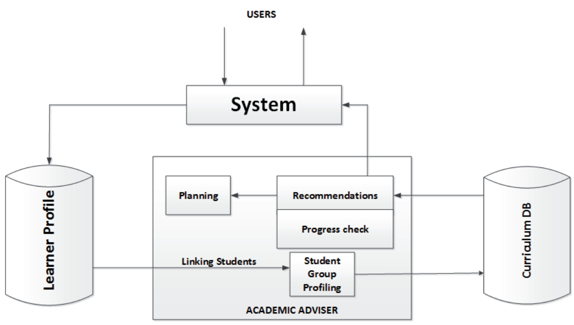


Figure 6: Architecture of the Academic Adviser.

2. Collection of information about the courses and completed studies.
3. Creation of study progress profiles along the lines of Section 3.
4. Linking the individual student to student peers with similar study profile in the institutional environment.
5. Student's progress check. If student is linked to a profile requiring intrusive advising, inform the adviser and the student by providing the interpreted study profile to support the communication and problem solving.
6. Modification of the study plan on recommended courses and their timetables by taking into account the evidence related to the identified study profile.
7. Planning and realization of an intrusive profile intervention adaptively for a larger pool of students.

Data collection related to the system is, naturally, all the time active. The evidence-based study profiles can and should be recomputed on regular basis. A natural suggestion would be to do this after the studies made during the previous semester have been stored and become available for clustering.

4.2 Automated Academic Advising process

The automated process of Academic Advising, related to the system's architecture and main functionality as described above, is presented in Figure 7. The given proposition allows manual control of the continuous advising activity for every learner individually, or the more automated process where the role of the advisers is shifted to the higher level of abstraction. The difference on the level of learner's life cycle between these two use cases is depicted in Figure 8. The automated process is highlighted with the red color in the figure. In the automated scenario, the responsible persons of the study organization only provide policies, planning and regulations. This can reduce the responsibility for the daily routine work and could help to provide recommendations for a larger pool of students rather than for the each individual learner.

The work-flow related to Figure 7 reads as follows: The learner is choosing the study program of an educational institution. After that he or she chooses with AA the proper courses which are related to the chosen program and creates a study plan. Information about the student, the required courses and progress in them is stored in the database and is automatically changed/refreshed after each passed course. After passing several courses, system can attach a student to a group of students with similar, actual study path. If learners are doing well, evidence-based determination and communication of this during advising encourages them to continue like that. If they are attached to a profile which does not progress with the studies as expected, the system can identify this early and provide intrusive academic advising support for both the advisor and the students in question.

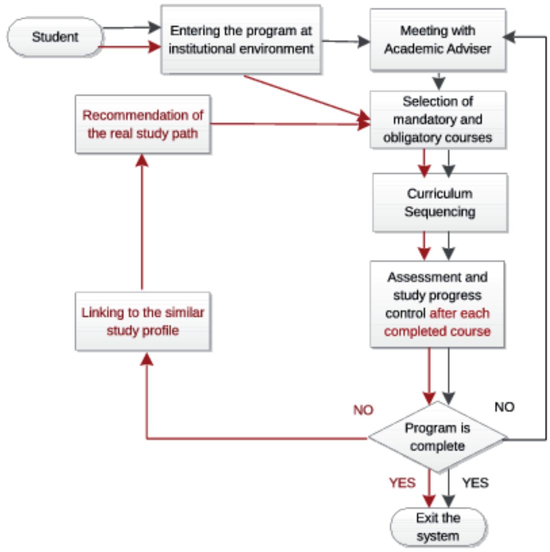


Figure 8: Comparison of the manual and automated processes of learner's study life cycle.

The proposed automated mechanism solve an important problem of improving and providing academic advising, because more and more students should receive guidance with their study plans before graduating. This system will help to plan when and how to provide the courses, especially the compulsory ones, as well as to plan a profile intervention adaptively for a larger pool of students, which will reduce the human effort of academic advising.

5 CONCLUSIONS

Academic Advising is an essential part of daily activities in an educational institution and an important component in the learner's study life. Nowadays, we need to be able to create and manage personalized study plans and study paths taking into account learners abilities and regulations of the learning environment. And in order to better help students, Academic Adviser should be able to manage a rich set of information, e.g., on short-range program planning, evaluation of students, and generation of the proper teach-

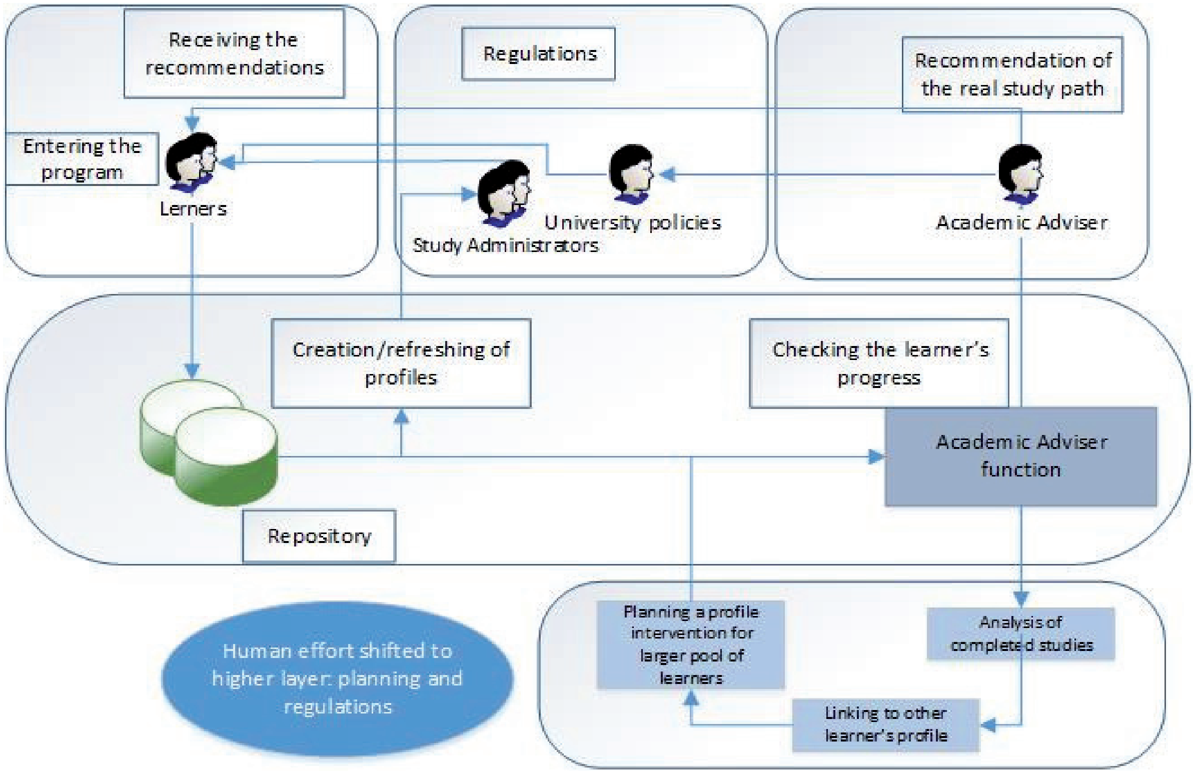


Figure 7: Automated Academic Advising process.

ing schedule, as well as plan possible interventions adaptively for a group of students instead of following all individual students separately. It is decisive that learner should receive proper advising – poor or no advising is known to have a negative effect on the progress in studies (Al-Ansari et al., 2015).

In this paper, we presented a compact literature review about Academic Advising, mostly focusing on Automated Academic Advising and Intelligent Academic Advising. It was then described how, by using a robust variant of prototype-based clustering method, which is especially suitable for data with missing values, one can create prototypical student group profiles characterizing the overall progress of the studies. This allows academic advisers to provide evidence-based information on the study paths that were actually realized by individual students. Moreover, academic institutions can focus on management and updates on course schedule having an effect on clearly characterized and recognized groups of students. Note that even if the sample groups of students that were profiled here were very small, the used method is scalable to hundreds of thousands of students (Saarela and Kärkkäinen, 2015b).

Then a reference model for automated Academic Advising system was proposed. The proposed architecture and model of the system are intended for a development phase to prototype the whole automated process, where the learners will be profiled regularly, and where the proper study path will be presented, as well as deviating learners detected. The proposed model of the AA system will have automated process of study path recommendation. This system will help to plan when and how to provide the courses, especially the compulsory ones, as well as to plan a profile intervention adaptively for a larger pool of students, which will reduce the human effort of academic advising.

By continuing the development of the line of work, we could consider the study paths with higher granularity than per semester. Also, the main functionality of the proposed system – to provide an automated notification for the academic advisers about students and their progress, with the interpretation of needs to modify and re-plan the study path – should be properly evaluated. Moreover, better availability of learner's personal information concerning the study entrance criteria and current life situation, e.g., a part-time job or living far from the institute, could support both interpretation of the generated student profiles and better preparation and management of the intervention patterns of academic advising.

REFERENCES

Al-Ansari, A., El Tantawi, M., AbdelSalam, M., and Al-Harbi, F. (2015). Academic advising and student support: Help-seeking behaviors among Saudi dental undergraduate students. *The Saudi dental journal*, 27(2):57–62.

Al-Sarem, M. (2015). Building a decision tree model for academic advising affairs based on the algorithm C 4-5. *arXiv preprint arXiv:1511.04026*.

Auvinen, T. (2015). *Educational Technologies for Supporting Self-Regulated Learning in Online Learning Environments*. Aalto University.

Auvinen, T., Paavola, J., and Hartikainen, J. (2014). STOPS: a graph-based study planning and curriculum development tool. In *Proceedings of the 14th Koli Calling International Conference on Computing Education Research*, pages 25–34. ACM.

Äyrämö, S. (2006). *Knowledge Mining Using Robust Clustering*, volume 63 of *Jyväskylä Studies in Computing*. University of Jyväskylä.

Biletskiy, Y., Brown, J. A., and Ranganathan, G. (2009). Information extraction from syllabi for academic e-Advising. *Expert Systems with Applications*, 36(3):4508–4516.

Brusilovsky, P. (1998). Adaptive educational systems on the world-wide-web: A review of available technologies. In *Proceedings of Workshop" WWW-Based Tutoring" at 4th International Conference on Intelligent Tutoring Systems (ITS'98), San Antonio, TX*.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., and Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5-6):318–331.

Chen, C.-M. (2008). Intelligent web-based learning system with personalized learning path guidance. *Computers & Education*, 51(2):787–814.

Chen, C.-M. (2009a). Ontology-based concept map for planning a personalised learning path. *British Journal of Educational Technology*, 40(6):1028–1058.

Chen, C.-M. (2009b). Personalized e-learning system with self-regulated learning assisted mechanisms for promoting learning performance. *Expert Systems with Applications*, 36(5):8816–8829.

Chen, C.-M. et al. (2008). Ontology-based concept map for planning personalized learning path. In *IEEE International Conferences on Cybernetics & Intelligent Systems*.

Chen, C.-M., Lee, H.-M., and Chen, Y.-H. (2005). Personalized e-learning system using item response theory. *Computers & Education*, 44(3):237–255.

Dougherty, S. B. (2007). Academic advising for high-achieving college students. *Higher Education in Review*, 4:63–82.

Drezner, Z. and Hamacher, H. W. (2001). *Facility location: applications and theory*. Springer Science & Business Media.

Drozd, D. S. (2010). *Student preferences for academic advisors as transformational leaders*. PhD thesis, Texas A&M University.

Feghali, T., Zbib, I., and Hallal, S. (2011). A web-based decision support tool for academic advising. *Educational Technology & Society*, 14(1):82–94.

Fung, A. and Yeung, J. (2000). An object model for a web-based adaptive educational system. In *Proceedings of the IFIP International Conference on Educational Use of Technologies (ICEUT2000)*.

Henderson, L. K. and Goodridge, W. (2015). AdviseMe: An intelligent web-based application for academic advising. *International Journal of Advanced Computer Science and Applications*, 6(8):233–243.

Huang, M.-J., Huang, H.-S., and Chen, M.-Y. (2007). Constructing a personalized e-learning system based on genetic algorithm and case-based reasoning approach. *Expert Systems with Applications*, 33(3):551–564.

Huber, P. J. (2011). *Robust statistics*. Springer.

Hwang, G.-J., Kuo, F.-R., Yin, P.-Y., and Chuang, K.-H. (2010). A heuristic algorithm for planning personalized learning paths for context-aware ubiquitous learning. *Computers & Education*, 54(2):404–415.

Idris, N., Yusof, N., Saad, P., et al. (2009). Adaptive course sequencing for personalization of learning path using neural network. *Int. J. Advance. Soft Comput. Appl*, 1(1):49–61.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, 31(8):651–666.

Kaihlavirta, A., Isomöttönen, V., and Kärkkäinen, T. (2015). A self-ethnographic investigation of continuing education program in engineering arising from economic structural change. *Studies in Continuing Education*, 37(1):99–114.

Latorre-Navarro, E. M. and Harris, J. G. (2015). An intelligent natural language conversational system for academic advising. *Editorial Preface*, 6(1).

Lee, M.-G. (2001). Profiling students' adaptation styles in web-based learning. *Computers & Education*, 36(2):121–132.

Leung, C. M., Tsang, E. Y., Lam, S., and Pang, D. C. (2010). Intelligent counseling system: A 24 x 7 academic advisor. *Educause Quarterly*, 33(4):n4.

Liu, C.-l., Wu, S., Chang, M., and Heh, J. (2008). Guiding students to do remedial learning in school campus with learning objects' spatial relations. In *16th international conference on computers in education conference*, pages 249–256. Citeseer.

Minaei-Bidgoli, B. (2004). *Data mining for a web-based educational system*. PhD thesis, Michigan State University.

Mostafa, L., Oately, G., Khalifa, N., and Rabie, W. (2014). A case based reasoning system for academic advising in egyptian educational institutions. In *2nd International Conference on Research in Science, Engineering and Technology (ICRSET2014) March*, pages 21–22.

Nguyen, T. B., Nguyen, D. N., Nguyen, H. S., Tran, H., and Hoang, T. A. D. (2008). An integrated approach for an academic advising system in adaptive credit-based learning environment. *Journal of Science, Natural Sciences and Technology*, (24):110–121.

Noaman, A. Y. and Ahmed, F. F. (2015). A new framework for e academic advising. *Procedia Computer Science*, 65:358–367.

Papanikolaou, K. A., Grigoriadou, M., Magoulas, G. D., and Kornilakis, H. (2002). Towards new forms of knowledge communication: the adaptive dimension of a web-based learning environment. *Computers & Education*, 39(4):333–360.

Saarela, M. and Kärkkäinen, T. (2014). Discovering Gender-Specific Knowledge from Finnish Basic Education using PISA Scale Indices. In *Proceedings of the 7th International Conference on Educational Data Mining*, pages 60–68.

Saarela, M. and Kärkkäinen, T. (2015a). Analysing Student Performance using Sparse Data of Core Bachelor Courses. *JEDM-Journal of Educational Data Mining*, 7(1):3–32.

Saarela, M. and Kärkkäinen, T. (2015b). Do Country Stereotypes Exist in PISA? A Clustering Approach for Large, Sparse, and Weighted Data. In *Proceedings of the 8th International Conference on Educational Data Mining*, pages 156–163.

Saarela, M. and Kärkkäinen, T. (2015). Weighted Clustering of Sparse Educational Data. In *Proceedings of the 23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN 2015)*, pages 337–342.

Saarela, M. and Kärkkäinen, T. (2017). Knowledge Discovery from the Programme for International Student Assessment. In Peña-Ayala, A., editor, *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning*, pages 229–267. Springer International Publishing, Cham.

Sandvig, J. and Burke, R. (2005). Aacorn: A cbr recommender for academic advising. Technical report, Technical Report TR05-015, DePaul University.

Schiaffino, S., Garcia, P., and Amandi, A. (2008). eteacher: Providing personalized assistance to e-learning students. *Computers & Education*, 51(4):1744–1754.

Sprent, P. and Smeeton, N. C. (2016). *Applied nonparametric statistical methods*. CRC Press.

Study Amanuensis (November 2, 2016). private communication.

Tang, T. Y. and McCalla, G. (2005). Smart recommendation for an evolving e-learning system: Architecture and experiment. *International Journal on elearning*, 4(1):105.

Tuttle, K. N. (2000). Academic advising. *New directions for higher education*, 2000(111):15–24.

Wallden, L. J. (2016). *Kansainvälisten koulutusarvioiden vertailu koulutuksellisen tiedonlouhinnan keinoin*. Jyväskylä Studies in Computing. University of Jyväskylä.

Wen, F. L. S. L. D. and McGreal, F. Z. K. R. (2015). e-Advisor: A multi-agent system for academic advising. *International Journal of Advanced Computer Science and Applications*, 6(8).

Werghi, N. and Kamoun, F. K. (2009). A decision-tree-based system for student academic advising and planning in information systems programmes. *International Journal of Business Information Systems*, 5(1):1–18.

Zaki, M. J. and Meira Jr, W. (2014). *Data Mining and Analysis: Fundamental Concepts and Algorithms*. Cambridge University Press.

PVI

SMART EDUCATIONAL PROCESS BASED ON LEARNING CAPABILITIES

by

Mariia Gavriushenko, Renny S.N. Lindberg, Oleksiy Khriyenko 2017

9th annual International Conference on Education and New Learning Technologies (EDULEARN 2017), Barcelona, Spain

SMART EDUCATIONAL PROCESS BASED ON PERSONAL LEARNING CAPABILITIES

Gavriushenko Mariia[1], Lindberg Renny S. N.[2], Khriyenko Oleksiy[1]

[1]*Faculty of Information Technology, University of Jyväskylä, Jyväskylä (FINLAND)*
[2]*Department of Software, Ajou University, Suwon (REPUBLIC OF KOREA)*

Abstract

Personalized learning is increasingly gaining popularity, especially with the development of information technology and modern educational resources for learning. Each person is individual and has different knowledge background, different kind of memory, different learning speed. Teacher can adapt learning course, learning instructions or learning material according to the majority of learners in class, but that means that learning process is not adapted to the personality of each individual learner. That is why it is important to have smart educational process based on personal learning capabilities. This paper presents a literature survey on different learning systems which detects learning progress and based on that a model of smart educational system which use knowledge engineering and Watson technology is proposed. This system is relevant both for basic education and for adult education.

Keywords: Smart educational process, personalisation, Ontology, Watson Technologies.

1 INTRODUCTION

Development of information technologies has quickly changed approach to learning and education, with the emergence of more mobile technologies and faster computers. A myriad of methods to deliver these educational innovations has emerged. For instance, with children educational process could be presented in a game format, or more social platform.

Developing e-learning has become more popular due to several benefits tied to it, such as: mobility and possibilities to customize user navigation [1], [2]. Modern methods of learning with the help of computers have tightly entered the educational process, from pre-school education to retraining. Generally the learning topics are presented in logical chain, moving from topic to topic. The process of enhancing the level of learning, in a company for instance, requires paying attention to the personal needs of the employees.

The problem of innovative teaching methods lies in their wide applicability due to the advantages mentioned above, but at the same time there are significant drawback – it is the passage to a new level with a specific threshold. When a person collects certain percentage of correct answers, he or she can proceed to the next level, not taking into account that part of the tasks were not performed or were performed incorrectly. This is an important issue, because most topics are interconnected and “gaps” in the learner’s knowledge should be avoided.

Also, an important issue is related to adult learning, when it is not necessary to start learning some topic from the scratch, but when person already has knowledge background in specific field and just needs to improve his knowledge in specific topic related to his background. This is especially significant for those people who need to learn more content and improve their qualification to be more competitive. Unfortunately, not many learning platforms take into account learner’s background while suggesting some material for additional learning and because of that learner has to spend more time for searching relevant material.

To solve these issues, we surveyed papers on existing learning systems to see how they deal with learner progress assessment. Based on our findings we designed our education support system. This system aims to make the process of learning more personalized taking into account assimilability of the content and existing knowledge background. Setting the threshold value of 100% is not a solution, however since the user might also lose interest, if the system is too demanding. Tracking the user's progress and general level of understanding is more important than superfluous obstacles.

2 BACKGROUND

A literature survey on e-learning was conducted to see how different systems detect users learning progress. 36 papers in total were accepted as part of our survey, which discuss: tutoring, learning or e-learning systems. Due to the fast pace of advancement of technological concepts and methods we would only accept papers that were published in 2010 or after. Google Scholar, ACM library and IEEE Xplorer were used to search for the publications.

Our core focus was on the handling of learning process and especially how the systems ensure that users understood the learning topic. Another point of interest was on how the system reacts to any detected learning difficulty. Out of the 36 papers discussing learning systems, 13 had no detection for possible difficulties that the user might be having. The remaining 23 systems approach to learning detection varied somewhat, but could be roughly labelled within two distinct types: By testing and user behaviour analysing.

The systems that used testing were labelled to three sub groups: loop, feedback and other. Systems that use loop commonly had either periodic quiz, or a larger questionnaire at the end of a learning topic and depending on the user's results the system would not proceed further before user learning level was at a satisfactory level. In total we labelled eight systems within the loop sub category. Depending on the user's answers the loop-test systems reacted in slightly varying ways, by either presenting more material [3], [4], [5] or simply show the same materials and not show any new materials before the system is satisfied that the user has reached a satisfactory learning level [6], [7], [8]. With two systems the systems would either return to the topic at later point [9] or only suggest the user to sticking with the current topic, but not enforcing it [10]

Learning systems labelled under the sub group "feedback", essentially these systems also have questionnaires that the user must complete in order to proceed further, but the system will only show results of the questionnaire before moving further. Only one learning system also presented links for additional learning material depending the user's results [11]. Only one learning system used the test results to modify internal features that in turn affect future quizzes [12].

Last sub group was simply labelled as other as they were slightly too varied to be placed under either loop or feedback. In the following Table 1 we show the evaluated learning systems.

The group labelled as "By analysis" simply put, utilize user behaviour and other various methods to try and deduce the level of knowledge by the user. These systems have essentially no testing of the user, however they do have profiles created that two of the systems updated by prompting for the user to give feedback at the end of a cycle [13], [14]

The remaining 13 surveyed systems had no learning level detection with one exception. One system essentially had no internal analysis of the learner's level of knowledge, but instead the system had the users assess and each other[15]

Table 1.List of surveyed papers on learning systems and what type of methods they use to detect user learning.

Method	Type	Paper				
By test	loop (8)	[3]	[6]	[7]	[5]	[9]
		[10]	[8]	[4]		
	feedback (6)	[16]	[17]	[18]	[11]	[19]
		[12]				
	other (4)	[20]	[21]	[22]	[23]	
By analysis (4)		[24]	[25]	[13]	[14]	
No detection (14)		[26]	[27]	[28]	[29]	[30]
		[31]	[32]	[33]	[34]	[35]
		[36]	[37]	[38]	[15]	

3 SMART EDUCATIONAL PROCESS

In this section we argue that there is a need for system which supports smart educational process based on personal learning capabilities. This process is important for both, basic education and adult education. Proposed method is taking into account person's knowledge background and rate of material compatibility.

3.1 Basic education process

For basic education –preschool, primary, secondary and higher– a person is usually expected at first to learn the education material given and then perform a certain set of tasks. The topics for study are sequential, from simple to more complex ones, which, in turn, can be based on several previous ones.

Unfortunately, majority of schools do not have time to ensure that all students have fully understood the underlying education materials during an ongoing course that form the crucial base for more advanced, but related topics. Thus there is a strong likelihood that students will form knowledge gaps that will only accumulate as the course progresses. At its worst this can lead to a student unable to proceed forward in the course and is thus forced to retake it.

The tasks (assignments) after the study of the material are presented in the form of a straight line. When a student solves several problems in a row correctly (assuming 10 tasks), this indicates that the student has thoroughly understood the topic and can proceed to next topic. If a student makes a mistake somewhere and answers incorrectly from the task cycle, then this test is followed by additional testing and the student cannot proceed to the next topic. An error is an incorrect answer or a long response rate (presumably, this time is spent searching information in external sources). Then, for this topic, and what is related to it, an additional set of tasks is given, which allows learner to fix the material that is not learned well enough. The student is again invited to solve the problem until he answers specific percentage of right answers correctly.

We introduce three amounts, which could be changed:

1) k – fixed amount of tasks;

2) $l = k - \%k$ – means how many mistakes person can make in k;

3) m – amount which experts put for correlation of true answers to false answers;

N_{true} – amount of correct answers;

N_{false} – amount of incorrect answers.

If this equation is correct, learner can move to the next level:

$$l \geq \frac{N_{true} \times \frac{k}{m}}{N_{false}}$$

The more mistakes a learner makes – the more time is spent by the learner on solving assignments and reading additional material until next level.

This approach allows for more dynamic learning and also minimizes the forming of any knowledge gaps. This approach also makes it easier for the system to deduct what specific topic the learner might be struggling with and when they are outside their zone of comfortable learning speed.

Each task should have its own weight (complexity). And complexity could be the sum of topics in the tasks (the weight of the tasks should be set by the experts or teachers). The severity of tasks correlates with the speed of the learner. If a person is slow to answer (by an average formed from all learners) this could indicate that one or more previous topics have not been mastered fully.

When a learner is stuck or makes mistakes, it is necessary to give them the proper guidance to help them. And if something is not clear from the explanation, then the person should have the ability to allocate a certain part of the material, which is incomprehensible to him and to receive a further explanation.

If a person quickly and successfully copes with the tasks, then the rate of learning should be optimized for the individual. For an adept learner this can be achieved with fewer assignments that are more complex that utilize several topics at once. It is not difficult for some students to quickly solve problems and move on to more complex topics and tasks that are based on several topics at once. If a person is capable of solving tasks without mistakes in a row, then he can be given less tasks on the next level and several subtopics to study and solve:

$$if \begin{cases} N_{true_i} = k_i, k_{i+1} = k_i - p \\ (N_{true_i} > k_i) \,\&\&(k_i < k_0), k_{i+1} = k_i + p \\ (N_{true_i} > k_i) \,\&\&(k_i \geq k_0), k_{i+1} = k_0 \end{cases}$$

If there is a combined task – then it is good to have a counter, which will take into account mistakes which the learner is making and where, as well as count which subtopic was answered incorrectly more times than others. After that system should provide additional material for study.

Because the system will have assignments that combine not only the current topics, but likely also additional sub topics the learned might have been struggling with. This can make the presented number of topics somewhat complicated, which necessities modifying the formula to mitigate this:

$$l \geq \frac{\sum_{i=1}^{N_{true}} \frac{d_{task\,level}}{d_{current\,level}}}{\sum_{i=1}^{N_{false}} \frac{d_{task\,level}}{d_{current\,level}}}$$

Tasks can include:

1) The choice of one of the proposed options for answers,

2) Formula or question, where the answer must be written in text form.

In order to analyse this kind of natural language input, there are novel cognitive computing approaches for the design. This means that communication between systems will occur more natural way.

There are several cognitive computing services that could be used for this analysis:

1) IBM Watson [39],

2) Google mind cognitive services like Microsoft Azure [40].

These services allow developing a natural language form.

3.2 Adult education process

The speed of constantly emerging new technologies, utilized by consumers and by the companies making these products for the consumers, is only gaining momentum. Because of that employees of large companies need to study constantly additional material, be aware of new products and other things in order to remain competitive and professional in their field.

This field of work as IT is an exponentially developing field of human activity, in which new tools, technology, applications are continuously emerging, thus it can be said that the IT sphere is where "all life learning" rule is applied.

Unlike a student who we assume starts from zero an employee of a technology company is likely to have some level of knowledge of a new topic he or she might be learning, with some knowledge gaps here and there. Our learning system would be able to target those gaps, before moving on the actual topic at hand.

If a person's knowledge is lower in a certain area, then this is a certain gap between the study material and his knowledge, and then knowledge retrieval is necessary, where a person reaches his level and then studies what is necessary. It should be taken into account that people have different knowledge base, so when searching for spaces, each person will have different top down (different depth in the material that needs to be studied). Some people should study the material superficially to sort out a certain topic, and some should go deeper in order to thoroughly understand all the nuances.

Unlike with a student, this form of “adult” education, does not force to the learner to approach the assignments of topics in set order. A person will learn from a case based study. The system will offer a practical example to study the material. If the system identifies a gap in the learners knowledge it will offer to present additional study materials in more detail. After the learner gets a detailed explanation, system searches in the database of the total content, provides a link for relevant material. This training is based on the principle of browsing.

4 SYSTEM'S MODEL PROPOSAL. METHOD AND ARCHITECTURE

An effective computer-supported learning tool that helps in learning process requires efficient knowledge representation, where it is possible to analyse data and make sense of it using specific rules.

Ontologies, which have proved themselves from ancient times as an approach for the integration of data and knowledge in science, are one of the modern methods allowing obtaining a quick idea of the content of an information resource. Ontologies are used for formal specification of concepts and their relations, which characterize a specific knowledge domain. For this system's model proposal, the advantage of using this approach of knowledge representation is a formal structure, which greatly simplifies the processing [41], [42], [29].

Ontologies have been used in educational domains to build, organize and update learning resources like learners profiles, learning objects, learning paths, academic performance, etc [43]. Ontology will be served as knowledge engineering process which will store and analyse data about learning topics, assignments, learning material and personal information of learners which shows there learning background. Every system which has knowledge engineering process has to have domain experts. These people will compose tasks and comment level of passing those tasks. Tasks could be written with different levels of complexity, which consists of one or several topics. Also, experts will provide information for each topic, also what will be an optimal minimum for checking if person learned topic or not. Each topic will have its own optimal minimum of tasks.

After learning some material learner has to pass specific amount of tasks and he has to solve them with specific threshold according to formulas presented in Section 3.1. The counter will be counting false answers and detecting knowledge gaps checking the relationships between topics in the Ontology.

In the proposed system, it is necessary to conduct an analysis of the learner's behavior. If system sees that a person makes mistakes, then after answers correctly, then again makes mistakes, then in this case it is necessary to conduct an analysis, and, perhaps, there is some pattern that the topics are related to other sub-themes, and in which percentage (some in less, some in more).

If person is doing well and he is fast in learning and in testing process, then on the next stage of testing he will get fewer tasks, but not less than defined minimum, for faster learning. But if person if making mistakes, then system is adding more assignments to solve which will consist also from the previous topics which are the basis for this particular one, and more learning material to study.

If learner makes a mistake in the task, he himself can allocate a specific piece of material that he does not understand. Then the system will offer him an explanation, based on which sub-themes served as the basis of this topic. The system will look how to fill these gaps. The system should be semi-automated and use technologies such as Watson Retrieve and Rank [44]. Retrieve and Rank service will search the most relevant information from a collection of documents. Using this service it is possible to collect and load material, train the machine learning rank model, query service and evaluate results. The person gets a topic to learn, does not understand the paragraph, singles it out and runs the service for more in-depth study of the material. For example, it is possible to find quickly some solutions from technical manual. If a person does not understand the piece of content, then he can dig deeper into the topic, where on the fly to find more relevant one, and dig deeper and deeper until he reaches his level of knowledge base with which he can start without problems to study the material.

Architecture of the proposed system model is presented on Fig.1.It shows main blocks of the smart educational process. Our approach follows very much the style that in the background section was labelled as "test by loop". Periodic testing is the only certain way to ensure that no large knowledge gaps emerge. Relying on just analysis or giving the users only feedback is not enough. We've built on the existing models slightly on how wrong answers affect the process. The user is not locked inside a small loop that concentrates on the topic at hand, rather the system will start pinpointing where the knowledge gap is. Thus the system also, in a manner, does a minor scale retention test for the user.

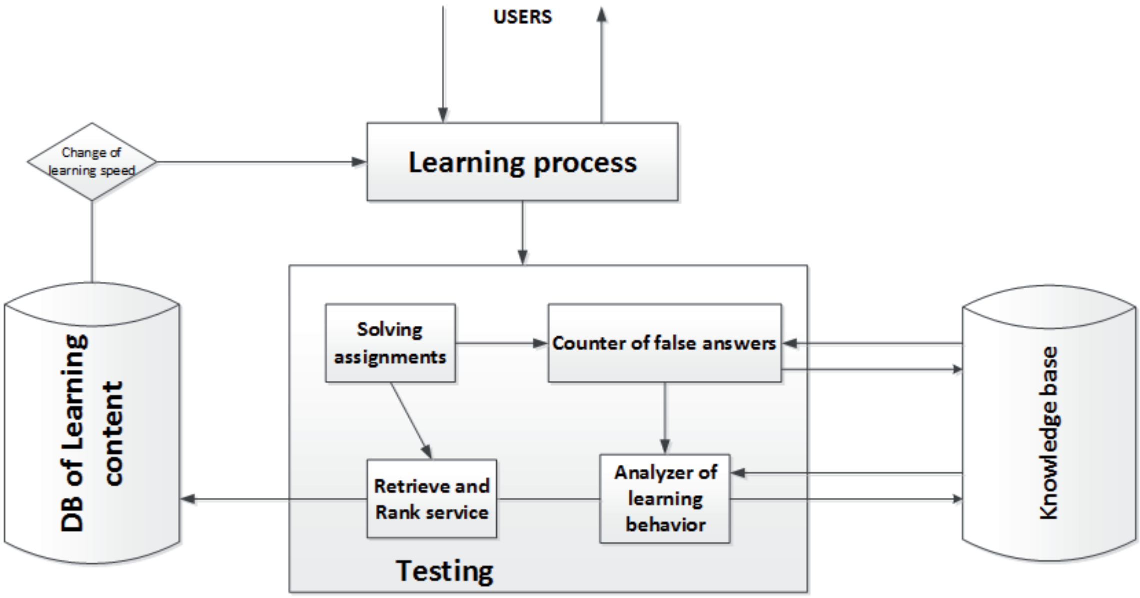


Figure 1. Architecture of smart educational process.

5 CONCLUSIONS

This paper propose solution for enhanced personalized learning system, which uses a technology of knowledge gaps reduction. It consists of the fact that at a certain threshold of passing threads, person will have the opportunity to proceed to the next level, but at the same time, the missing answers will be taken into account at the next level. System will analyse where learner is making mistakes and find closely related topics which he or she has to study more. This system is an Ontology-driven system, because Ontology is a good instrument for presenting a knowledge base of learning process and relationships between different topics, as well as for creating a new knowledge. This system builds on the several surveyed systems and aims to utilize and enhance on their existing methods, such as internal analysis and personalized testing.

Particular solution allows personalizing learner's educational gaps due to the fact that incorrect answers will be taken into account in the next level. Accordingly, the student, who will take the whole course due to the rules, will be fully familiar with the material, which is proposed for the study. Also it will help on-the-fly in improving adult's knowledge in some particular domain. This technique also allows future maintaining of performance records according to various criteria, for example: to determine what topic is the most incomprehensible, also possible to draw conclusions about what is the relationship between age and the assimilation of the material, as well as any analytical conclusion is based on statistics, which in future will successfully affect the tuning (change) of the learning system.

6 ACKNOWLEDGMENTS

We would like to give our thanks to Aziz Hasanov for his invaluable help with the learning system survey during the writing of this paper.

REFERENCES

[1] Garrison, D. Randy. E-learning in the 21st century: A framework for research and practice. Taylor & Francis, 2011.

[2] K. Mac Keogh and S. Fox, "Strategies for Embedding eLearning in Traditional Universities: Drivers and Barriers," vol. 7, no. 2, pp. 135–141, 2008.

[3] A. Latham, K. Crockett, D. McLean, and B. Edmonds, "A conversational intelligent tutoring system to automatically predict learning styles," *Comput. Educ.*, vol. 59, no. 1, pp. 95–109, 2012.

[4] G. J. Hwang, H. C. Chu, J. L. Shih, S. H. Huang, and C. C. Tsai, “A decision-tree-oriented guidance mechanism for conducting nature science observation activities in a context-aware ubiquitous learning environment,” *Educ. Technol. Soc.*, vol. 13, no. 2, pp. 53–64, 2010.

[5] C. R. Beal, “AnimalWatch: An Intelligent Tutoring System for Algebra Readiness,” in *International Handbook of Metacognition and Learning Technologies*, vol. 28, no. January, R. Azevedo and V. Aleven, Eds. Springer, 2013, pp. 337–348.

[6] M. Siadaty *et al.*, “Learn-B: A Social Analytics-enabled Tool for Self-regulated Workplace Learning,” *Proc. 2nd Int. Conf. Learn. Anal. Knowl. - LAK ’12*, p. 115, 2012.

[7] P.-H. Wu, G.-J. Hwang, L.-H. Su, and Y. Haung, “A Context-Aware Mobile Learning System for Supportive Cognitive Apprenticeships in Nursing Skills Training,” *Educ. Technol. Soc.*, vol. 5, no. 1, pp. 223–236, 2012.

[8] Ö. Özyurt, H. Özyurt, A. Baki, and B. Güven, “Integration into mathematics classrooms of an adaptive and intelligent individualized e-learning environment: Implementation and evaluation of UZWEBMAT,” *Comput. Human Behav.*, vol. 29, no. 3, pp. 726–738, 2013.

[9] I. H. Hsiao, S. Sosnovsky, and P. Brusilovsky, “Guiding students to the right questions: Adaptive navigation support in an E-Learning system for Java programming,” *J. Comput. Assist. Learn.*, vol. 26, no. 4, pp. 270–283, 2010.

[10] F. Bouchet, J. M. Harley, G. J. Trevors, and R. Azevedo, “Clustering and Profiling Students According to their Interactions with an Intelligent Tutoring System Fostering Self-Regulated Learning,” *JEDM - J. Educ. Data Min.*, vol. 5, no. 1, pp. 104–146, 2013.

[11] C. Limongelli, F. Sciarrone, M. Temperini, and G. Vaste, “The Lecomps5 framework for personalized web-based learning: A teacher’s satisfaction perspective,” *Comput. Human Behav.*, vol. 27, no. 4, pp. 1310–1320, 2011.

[12] V. A. Nguyen, V. C. Pham, and S. D. Ho, “A Context-Aware Mobile Learning adaptive System for Supporting Foreigner Learning English,” *Comput. Commun. Technol. Res. Innov. Vis. Futur. (RIVF), 2010 IEEE RIVF Int. Conf.*, pp. 1–6, 2010.

[13] B. Zhang, C. Yin, B. David, Z. Xiong, and W. Niu, “Facilitating professionals’ work-based learning with context-aware mobile system,” *Sci. Comput. Program.*, vol. 129, pp. 3–19, 2016.

[14] D. Gallego, E. Barra, S. Aguirre, and G. Huecas, “A model for generating proactive context-aware recommendations in e-Learning systems,” in *Proceedings - Frontiers in Education Conference, FIE*, 2012.

[15] S. Gómez, P. Zervas, D. G. Sampson, and R. Fabregat, “Context-aware adaptive and personalized mobile learning delivery supported by UoLmP,” *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 26, no. 1, pp. 47–61, 2014.

[16] B. Vesin, M. Ivanović, A. Klašnja-Milićević, and Z. Budimac, “Protus 2.0: Ontology-based semantic recommendation in programming tutoring system,” *Expert Syst. Appl.*, vol. 39, no. 15, pp. 12229–12246, 2012.

[17] I. Arroyo, B. P. Woolf, J. M. Royer, M. Tai, and S. English, “Improving math learning through intelligent tutoring and basic skills training,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6094 LNCS, no. PART 1, pp. 423–432, 2010.

[18] S. D’Mello, A. Olney, C. Williams, and P. Hays, “Gaze tutor: A gaze-reactive intelligent tutoring system,” *Int. J. Hum. Comput. Stud.*, vol. 70, no. 5, pp. 377–398, 2012.

[19] A. Harchay, L. Cheniti-Belcadhi, and R. Braham, “A Context-Aware Framework to Provide Personalized Mobile Assessment,” *Interact. Des. Archit.*, pp. 82–97, 2014.

[20] R. Z. Cabada, M. L. Barrón Estrada, and C. A. Reyes García, “EDUCA: A web 2.0 authoring tool for developing adaptive and intelligent tutoring systems using a Kohonen network,” *Expert Syst. Appl.*, vol. 38, no. 8, pp. 9522–9529, 2011.

[21] J. E. Gomez, J. F. Huete, and V. L. Hernandez, “A Contextualized System for Supporting Active Learning,” *IEEE Trans. Learn. Technol.*, vol. 9, no. 2, pp. 196–202, 2016.

[22] C.-C. Chen and P.-H. Lin, “Development and evaluation of a context-aware ubiquitous learning environment for astronomy education,” *Interact. Learn. Environ.*, vol. 24, no. 3, pp. 644–661,

2016.

[23] M. Li, H. Ogata, B. Hou, N. Uosaki, and Y. Yano, “Personalization in context-aware ubiquitous learning-log system,” *Proc. 2012 17th IEEE Int. Conf. Wireless, Mob. Ubiquitous Technol. Educ. WMUTE 2012*, pp. 41–48, 2012.

[24] R. Benlamri and X. Zhang, “Context-aware recommender for mobile learners,” *Human-centric Comput. Inf. Sci.*, vol. 4, no. 1, p. 12, 2014.

[25] Z. Sevarac, V. Devedzic, and J. Jovanovic, “Adaptive neuro-fuzzy pedagogical recommender,” *Expert Syst. Appl.*, vol. 39, no. 10, pp. 9797–9806, 2012.

[26] M. Vladoiu and Z. Constantinescu, “U-Learning within a Context-Aware Multiagent Environment,” *Int. J. Comput. Networks Commun.*, vol. 3, no. 1, pp. 1–15, 2011.

[27] E. Popescu, “Adaptation provisioning with respect to learning styles in a Web-based educational system: An experimental study,” *J. Comput. Assist. Learn.*, vol. 26, no. 4, pp. 243–257, 2010.

[28] S. Alkhuraiji, B. Cheetham, and O. Bamasak, “Dynamic Adaptive Mechanism in Learning Management System Based on Learning Styles,” *2011 IEEE 11th Int. Conf. Adv. Learn. Technol.*, pp. 215–217, 2011.

[29] N. Capuano, M. Gaeta, S. Salerno, and G. R. Mangione, “An Ontology-Based Approach for Context-Aware e-Learning,” in *2011 Third International Conference on Intelligent Networking and Collaborative Systems*, 2011, pp. 789–794.

[30] S. L. Wang and C. Y. Wu, “Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system,” *Expert Syst. Appl.*, vol. 38, no. 9, pp. 10831–10838, 2011.

[31] M. Yaghmaie and A. Bahreininejad, “A context-aware adaptive learning system using agents,” *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3280–3286, 2011.

[32] N. Kasaki, S. Kurabayashi, and Y. Kiyoki, “A geo-location context-aware mobile learning system with adaptive correlation computing methods,” *Procedia Comput. Sci.*, vol. 10, pp. 593–600, 2012.

[33] H. O. Chorfi, A. Z. Sevkli, and F. Bousbahi, “Mobile learning adaption through a device based reasoning,” *Cyprus Int. Conf. Educ. Res.*, vol. 47, pp. 1707–1712, 2012.

[34] K. Scott and R. Benlamri, “Context-Aware Services for Smart Learning Spaces,” vol. 3, no. 3, pp. 214–227, 2010.

[35] T.-C. Yang, G.-J. Hwang, and S. J.-H. Yang, “Development of an adaptive learning system with multiple perspectives based on students’ learning styles and cognitive styles,” *Educ. Technol. Soc.*, vol. 16, no. 4, pp. 185–200, 2013.

[36] M. Yasir and S. Sharif, “An approach to Adaptive E-Learning Hypermedia System based on Learning Styles (AEHS-LS): Implementation and evaluation,” *Int. J. Libr. Inf. Sci.*, vol. 3, no. January, pp. 15–28, 2011.

[37] F. Akbari and F. Taghiyareh, “E-SoRS: A personalized and social recommender service for E-learning environments,” in *8th National and 5th International Conference on e-Learning and e-Teaching, ICeLeT 2014*, 2014, pp. 1–12.

[38] M. Salehi, I. Nakhai Kamalabadi, and M. B. Ghaznavi Ghoushchi, “An effective recommendation framework for personal learning environments using a learner preference tree and a GA,” *IEEE Trans. Learn. Technol.*, vol. 6, no. 4, pp. 350–363, 2013.

[39] “IBM Watson - Build Your Cognitive Business with IBM.” .

[40] “Azure Cognitive Services Documentation - Tutorials, API Reference | Microsoft Docs.” .

[41] L. Zhou, “Ontology learning: state of the art and open issues,” *Inf. Technol. Manag.*, vol. 8, no. 3, pp. 241–252, Aug. 2007.

[42] M. Gavriushenko, M. Kankaanranta, and P. Neittaanmaki, “Semantically enhanced decision support for learning management systems,” in *Semantic Computing (ICSC), 2015 IEEE International Conference on*, 2015, pp. 298–305.

[43] M. Gaeta, F. Orciuoli, and P. Ritrovato, “Advanced ontology management system for personalised e-Learning,” *Knowledge-Based Syst.*, vol. 22, no. 4, pp. 292–301, 2009.

[44] “Retrieve and Rank | Overview of the Retrieve and Rank service | IBM Watson Developer Cloud.” .

PVII

ON DEVELOPING ADAPTIVE VOCABULARY LEARNING GAME FOR CHILDREN WITH AN EARLY LANGUAGE DELAY

by

Mariia Gavriushenko, Iida Porokuokka, Oleksiy Khriyenko, Tommi Kärkkäinen
2017

Manuscript for GSTF Journal on Computing (JoC)

On developing adaptive vocabulary learning game for children with an early language delay

M. Gavriushenko., I.Porokuokka, T. Kärkkäinen, O.Khriyenko

Abstract—**Vocabulary replenishment is an ordinary child development process. Deviations in this process can significantly affect the further progress and perception of the educational material in the school. A hypothesis was proposed that the similarity of words on various factors can influence the child's understanding. To test this hypothesis in this work, we propose the development of AdapTalk game for children. This game is concentrated on learning words, in the context of animals. The game will create basis for further testing the influence of the semantic similarity of words for the child. This work describes the development background, the basic principles of calculating the semantic similarity of words, as well as the main game processes, such as filtering, training, and testing. The proposed functionality is a good base for introducing other factors of similarity of words, further testing and introduction of an adaptive mechanism for improving the vocabulary of children.**

Keywords-adaptive learning, semantic distance, learning game, late talkers.

I. Introduction

Late talkers are children who start learning words and forming sentences later than their typically developing peers who start putting words together before the age of 2 (Preston et al., 2010). While many of these children later catch up with their peers, the delay in the early development of the reading skill has been connected with later difficulties with language learning, such as dyslexia (Lyytinen et al., 2001; Lyytinen, 2015; Lyytinen et al., 2005). A recent finding (Lyytinen, 2015) confirmed earlier observation (Lyytinen et al., 2005) revealing that if the delay considers receptive language, i.e. comprehension of spoken language during early years, such children will face serious difficulties in becoming fully literate.

A lot of attention has been given to providing interventions for these children to avoid the difficulties in the future. For example, Marulis and Neuman (2013) reviewed 67 studies on young children's oral language development. Positive general effects on the impact of vocabulary interventions in the early years were reported, but their efficiency was suggested to be improved to succeed in closing the gap in preschool aged children. Continuation meta-analysis by the same authors (Marulis and Neuman, 2013) concluded similarly emphasizing the need of creating powerful enough interventions, especially for children with low socio-economical status.

Games have been widely recognized to generate a positive effect on learning with motivating and engaging, which they do most successfully when they succeed in facilitating a flow experience (Kiili 2005). Flow is a state of mind where a surroundings and track of time is lost while being fully occupied with doing tasks that are of optimal difficulty level - not difficult which could cause anxiety and not too easy either, since that could lead to boredom (Csikszentmihalyi, 1990). Learning with mobile devices is recognized to be motivational as well as sufficient in keeping children's attention (Skiada et. al. 2014). Facilitation with this technology provides a cost-efficient basis for developing learning environments that are able to adapt tasks to optimal difficulty for each learner individually.

In our previous work, we described the general idea of an adaptive gamified word-learning application for late talkers (AdapTalk), whose adaptivity is based on the similarity distance mapping (Gavriushenko et. al. 2015). The software is to be used by children of age one and up, and teaching of words is done with simple multiple-choice tasks. These tasks include a sound heard from headphones and an array of pictures on a screen of a mobile device of which the learner then tries to choose the correct one. The role of the adaptivity algorithm in this case is to choose an appropriate array of pictures. In our previous work, we stated a hypothesis that the tasks would be easier if the array of other pictures than the correct one would be different according to a chosen similarity mapping, and vice versa, that the tasks would be harder if the array was chosen to include the most similar ones. Testing this hypothesis provides its own challenge for both developing the software and experiment design, and in this paper we concentrate on describing these aspects.

This paper consists of description of adaptive learning for late talkers (Chapter 2); proposed system description (Chapter 3); semantic similarity calculation methods and method that was used in particular work (Chapter 4); game process description itself, which include filtering, learning and testing (Chapter 5); and conclusions with description of future work (Chapter 6).

II. Adaptive Learning For Late Talkers

A. Late Talkers

Diagnosis of late talkers can be done through estimating the child's vocabulary size and ability to form word pairs, while ruling out other developmental deficits. Many late talkers catch up with their typically developed peers sooner or later, but some do not achieve age-appropriate language skills even as adults (Leonard, 2014). In later years these children can be diagnosed with Dyslexia (Lyytinen, 2004) and Specific Language Impairment (SLI). SLI can be diagnosed during the pre-school years (age three to five) and means a deficit specific to language that is not related to hearing loss, neurological damage, and nonverbal intelligence, and it is expected to persist (Leonard 2014, p. 3-4).

Some research suggests that LT's word learning strategies are rather a-typical than delayed. Beckage, Smith and Hills (2010, 2011) analyzed LT's and typically developing childrens' vocabularies semantic connections. Authors' method included retrieving the vocabularies with Bates-MacArthur word form (Fenson et. al. 1993) and making the connections to the collected words from co-occurrences from CHILDES, a dictionary of child-directed speech. Both studies concluded that LT's network structures were less connected – containing words with less frequent connections. Authors presented an explanation that they called the 'odd-ball-model' where the child would rather learn a word that is less connected to other words as if that would be more interesting. This model is in line with Borovsky, Schafer and Fahey's (2016) eye-tracking study's results where authors found that toddlers would answer multiple-choice questions incorrectly more likely when the distracting item was previously unknown. Both results suggest that LT's would be picking their interest from rather unknown than known words on contrary to word-learning studies of adults and typical learners where, as Borovsky (2015) summarizes, the results commonly suggest that a word is more likely to enter the vocabulary if it shares connections to other previously known words. Borovsky et. al. (2016) however did not find differences in network structures when vocabulary size was controlled, while Beckage et. al. (2011) did.

If LT's learning strategies are indeed more a-typical than just delayed, developing an adaptive learning environment is a relevant approach to address the need of developing more effective interventions.

B. Adaptive word learning

It is a well-known fact that personalized approach is beneficial for any learning process (Murray and Pérez, 2015). This knowledge and background from educational theories should be used while creating technologically improved learning environments that adjust to the student's needs. These learning environments are often referred to as "adaptive" or "intelligent" studying systems.

In accordance with (Gifford, 2013), adaptive learning is a system of methods focused on "creating a unique learning experience for each person through the introduction of computer software". Adaptive learning systems allow organizing content, recognizing the method to learn based on student's knowledge, and use individual estimation results to ensure personalised feedback for each learner (Sonwalkar, 2005).

Adaptive learning systems have a lot of attributes and functions (Venable, 2011) brought together to produce appropriate content, assistance and to lead the learner through the adaptive learning tasks or units: pre-test, pacing and control, feedback and estimation, progress tracking and reports, motivation and reward.

These systems can be uncomplicated or algorithm-based (Oxman et al., 2014). The more basic ones are based on rules, established with the usage of if-then-statements. Algorithm-based systems apply advanced mathematical formulas and machine learning ideas to adjust to individual learners more specifically. Earlier research by (Brusilovsky and Peylo, 2003) divides these systems into Adaptive Hypermedia Systems and Intelligent Tutoring Systems. These technologies can be used for adding adaptive or intelligent functions into learning systems.

Adaptive Hypermedia Systems involve adaptive presentation and adaptive navigation assistance, and also adaptive information filtering, which includes collaborative filtering and content-based filtering. Intelligent Tutoring Systems consist of intelligent solution research, curriculum sequencing, problem solving guidance and intelligent collaborative learning, which includes adaptive group formation and adaptive collaborative assistance.

Discovering an optimum way to introduce the notions to the children during their first years of life might benefit in reducing their risk to have troubles in language learning in after years. That is the reason for us to have adaptive content presentation in learning system for late talkers. We should use the adaptive presentation technology to make it accessible. The goal is to adjust the content of a hypermedia page to the user's aims with knowledge stored in the user model (Gavriushenko et al., 2017). In an adaptive presentation system, the pages are dynamically generated or built from the pieces for every user (Brusilovsky, 1999). The content presented to the kids should not be predetermined, but should have the power to adapt in accordance with user's responses.

We suppose that recognizing some special notion among several other similar ones is a more difficult task. Every child is able to embrace pictures and sounds corresponding to them in different ways. For one child, it is difficult to differentiate between words, which sound alike. For another child, it could be difficult to differentiate between words that have a similar pattern, whether they could have a similar shape or similar color. Another child may have difficulty with distinguishing words, which are semantically close to each other.

III. System Description

Empirical evidence in the learning game environment for language development and disorders, as documented e.g. in (Ronimus et al., 2014), emphasizes the role of perceptual differentiation of acoustically close phonemes, i.e., acoustic dissimilarity. Here we generalize this starting point by hypothesizing that concept learning and identification depend on personal comprehension of multiple parameters, such as sound and optical representations, as well as their semantics. Our method is based on processing the complexity level of concept identification induced by these factors. Taking accuracy of user's' replies into consideration, the proposed system will automatically increase or reduce the level of complexity, adjusting it to the personal learning abilities of the users, and in such a way will improve learning process.

The functionality of the approach, presented in the previous paper (Gavriushenko et al., 2016) takes into account differences between visual and auditory stimuli. Visual similarity for different concepts could be measured using Euclidean distance because of its simplicity (Wang et al., 2005), or with Tangent distance (Simard et al., 1993), etc. Phonetic similarity could be measured using string-based techniques like Edit Distance, Jaro-Winkler measure (Winkler, 1999), N-gram similarity function (Kondrak, 2005), Soundex Mokotoff (Mokotoff, 1997). More detailed information about visual and phonetic similarity measures are given in (Gavriushenko et al., 2016).

Assuming that sets of more alike ideas cause more difficulties for their identification (recognition), we should automate the process of picture selection using multidimensional concept resemblance metric, measured as a collection of similarity values of the three mentioned factors. After, it is possible to personalized proposed learning instrument via modifying the delta (Δ) of motion similarity and "fiddling" with various levels of impact of the three factors on collective resemblance.

$$ConceptSim\ (c_i, c_j) = F(Sim_V, Sim_P, Sim_S) \quad (1)$$

Since all of our concept similarity factors (visual, phonetic and semantic) could be represented as weighted functions of various similarity measuring techniques, we defined a general formula to calculate similarity between the concepts in (Gavriushenko et.al., 2016).

$$ConceptSim\ (c_i, c_j) = w_V * \sum_{k_V=0}^{n_V} w_{k_V} * Sim_{V_{k_V}} + w_P * \sum_{k_P=0}^{n_P} w_{k_P} * Sim_{P_{k_P}} + w_S * \sum_{k_S=0}^{n_S} w_{k_S} * Sim_{S_{k_S}} \quad (2)$$

where w_V, w_P, w_S are coefficients of weighted influence of visual, phonetical and semantic factors on concept similarity; n_V, n_P, n_S are number of different similarity measuring techniques used to calculate visual (Sim_V), phonetical (Sim_P) and semantic (Sim_S) similarity with corresponding weights of their influence ($w_{k_V}, w_{k_P}, w_{k_S}$).

General architecture of adaptive vocabulary learning was also presented in (Gavriushenko et al., 2016). It was concentrated on the functionality that if the learners improve, then the system should be able to provide them with more challenging tasks.

In this particular work, we concentrate more on semantic similarity, not on other ones, for future testing of our hypothesis that semantic similarity may influence learner's speed of learning, as well as their ability to differ words.

IV. Semantic Similarity

Determination of semantic similarity between ideas, i.e. comparable objects like different words or texts, has a central role in semantic sense comprehension. To measure similarity is not a straightforward task. The most widespread method is to measure semantic resemblance based on domain ontology; a theoretical model that depicts the appropriate domain.

B. Ontology-based Semantic Similarity

II. Ontology-based semantic similarity can be estimated with various methods (Sanchez, 2012; Bin et al., 2009):

a) Edge-counting method (graph-based method): computes the minimum path length connecting the appropriate ontological nodes through the 'is-a' -links (Sanchez et al., 2012). Identically remote pairs of ideas on the upper level of taxonomy are computed less similar to the ones, which refer to a lower level (Wu and Palmer, 1994). Frequently the depth of ontology is united with the shortest path length in a nonlinear function (Li et al., 2003), sometimes with the overlap between the nodes (Alvarez and Lim, 2007). In (Al-Mubaid and Nguyen, 2006), authors applied cluster-based measure on top of a minimum length path and taxonomical depth. Taxonomical edge-counting technique was extended by involving non-taxonomic semantic links to the idea in the path (Hirst and St-Onge, 1998).

b) Feature-based measures: utilize taxonomical features retrieved from ontology. The resemblance between two notions can be determined as a function of their general and differential attributes (assessing similarity between ideas as a function of their attributes) (Sanchez et al., 2012). Such facet-based classification could be united with similarity of common properties' values.

c) Combined measures: Combined measures which involve the edge-calculation based on information content (IC)-based measures with edge weight (Jiang and Conrath, 1997).

A. Semantic Classification

In our current solution, we have limited our focus to a domain of animals. Each domain brings certain specifics to the metric measuring semantic similarity. Semantic similarity can be measured differently depending on the context. The context dependent similarities could be calculated separately and be further aggregated using different weights for different contexts. Thus, for the chosen domain, we may highlight several classifications:

- Geolocation-based classification: here we distinguished animals by geographical regions they live in. It is a complex metric that integrates continent-based, latitude- and climate zone based clustering. Similarity between the clusters is calculated based on climate groups' hierarchy and similarity of different continents.
- Domestication-based classification: animals could be also divided to those who are fully domesticated by human and live at their homes and farms, those whom we may meet in a zoo, and those who live only in wild nature and most probably are only seen via video records and photos. In this case, distances between the

classes could be predefined. Also, other metrics that define semantic similarity in other contexts may exist. Therefore, analogically to other similarity factors, final semantic similarity measure could be aggregated by weighted products/sum.

- Biological species-based classification: this metric is based on subclass hierarchy of animals classified by biological families of animals. In this case, the most suitable approaches to measure similarity are graph-based techniques (e.g. Jaccard metric, Scaled shortest path, Depth of the subsumer and closeness to the concepts, etc.) (Euzenat and Shvaiko, 2013; Bouquet et al., 2004; Leacock and Chodorow, 1998; Haase e. al., 2004). Thus, we calculate semantic similarity of concepts based on locations of corresponding nodes in the graph, using taxonomy-based ontology that represents class hierarchy of animals.

B. Semantic Similarity for Particular Case

The semantic similarity is calculated as distance between the nodes in a semantic network. Our network is in fact a biological species-based classification. The network was created with webprotege (http://webprotege.stanford.edu/) from a basis found from the same platform. The ontology was then modified by deleting classes, which do not have an instance in our collection of words and further by stabilizing classes and their child nodes so that each animal would be equally far from the center node. Fig.1 presents visualization of ontology of animals. On Fig.2 part of Coldblooded animals are presented. On Fig.3 sub taxonomy of Warmblood animals is depicted. The categorizations nature has the advantage of mapping more similar *looking* animals closer to each other. For example, dog and bear which are both mammals will be counted as more similar, while a semantic map formed from e.g. co-occurrences from children's books (e.g. fish and chicken are both frequently served dishes) might have disadvantages of mapping similarities between visual forms.

V. Process Description

The aim of this work is to test the hypothesis, that semantic distance influences children's' learning, as well as to see how children behave using this system. The system is a game-based learning environment for mobile phones and tablets, which presents tasks to a learner where one should point a referent from an array of pictures according to a sound (the object's name) which is heard from headphones.

The main goal is to determine which distractors affect learning more, choosing the pictures from the ones that are more similar versus less similar to the target. More precisely, whether choosing the right referent item is easier if the distractors are not similar to the target item.

To determine this, we chose to provide two main settings for the application. One that provides trials with most similar items and one with most distant items. To minimize the effect of previously known words and to determine results, we divided the game into three phases: filtering, learning and testing.

The filtering phase provides two trials per each item until the learner has answered incorrectly to target amount of items. Two trials per item were chosen, because answering twice correctly by chance is less likely than doing it once.

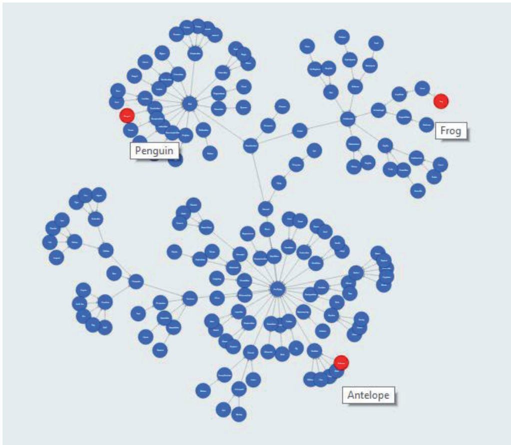


Figure 1. Vizualization of Ontology from WebVOWL (http://visualdataweb.de/webvowl/)

Once enough previously unknown items are collected, these items are then included in the learning phase, which provides four trials for each item. The first trial of learning phase presents the item alone and expects the learner to choose it. This is immediately followed by two trials with different distractors.

When all the learning items have been presented in four trials, the system proceeds to the testing phase. In this phase, all target items are presented once more with the same distractors as in the filtering phases for the first trial.

The distractors are chosen as follows: the semantic distance matrix is sorted for each concept starting from the most semantically distant (means that distant between words is minimal) to more semantically close (means that distance between words is minimal). Then we will be showing the most semantically distant distractor concepts for avoiding semantical influence on the process of filtration, learning and testing.

The amount of trials, which should be presented, in one level is calculated from approximates as follows. One trial consists of (1) presenting the trial and the sound, (2) waiting for the answer from player and (3) the feedback animations, each taking approximately two seconds. In sum, this would be six seconds per trial, which means that we can fit approximately ten trials in one minute. In this case, we chose to fit the levels with ten trials in filtering, nine in learning and ten in filtering. Like this, one filtering level will contain five different items, a testing level contains ten, and a single learning level fits handling three items. This means that ten-minute exposure time for trials is enough for 90 trials. The time between levels is not controlled, and not calculated as stimulation time.

Figure 2. CLASSES OF COLDBLOOD ANIMALS

C. Filtering Process

In the first phase, the aim is to filter out the words which Filtering Process the learner already knows, so that we will not include them as words which the child should learn during playing. Preliminary test at the beginning of the game checks if child knows any concepts from the predefined pool of the concepts. For example, child receives 20 pictures of concepts and the system checks each of the concept. If the child answers correctly about some particular concept few times, then system reduce this concept from the pool of presented concepts and until there will be only set of "unknown" concepts for that specific child.

Filtering process steps are as follows (Fig. 4):

1. We are building a matrix of distances (semantic distances). This matrix will show how particular concepts are distant (how they are not similar) to other concepts.

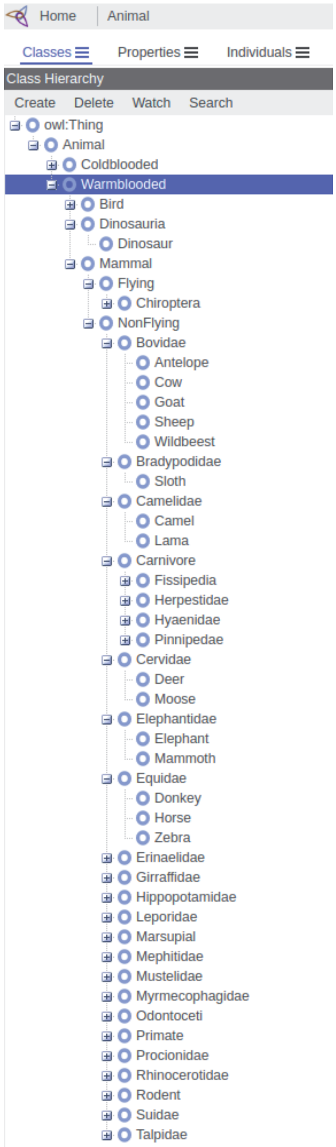


Figure 3. CLASSES OF WARMBLOOD ANIMALS

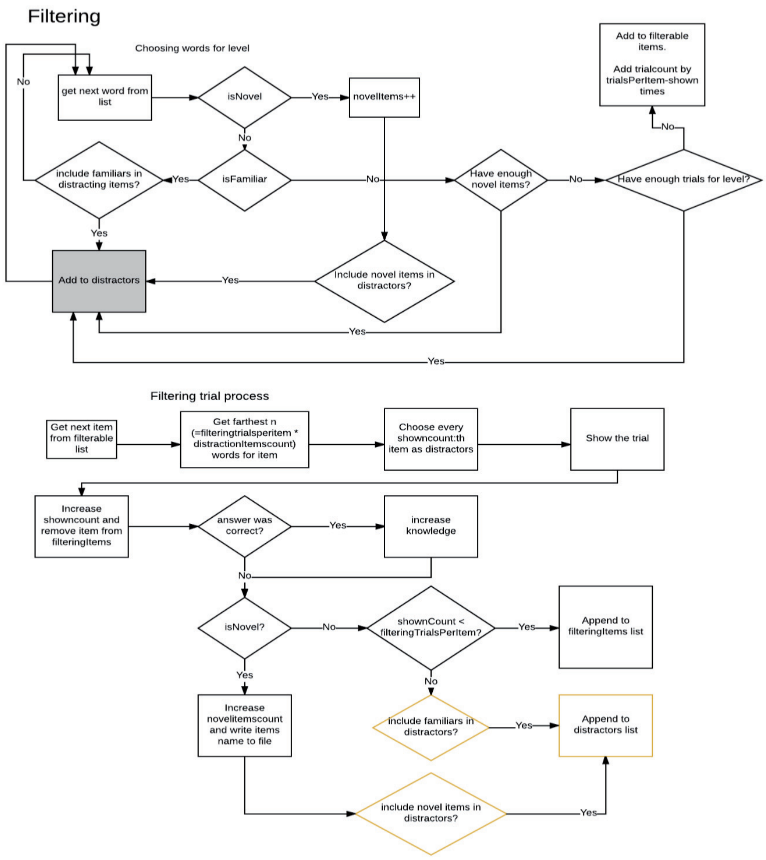


Figure 4. FILTERING PROCESS

2. We are sorting matrix for each concept from the most semantically distant (means that distance is minimum) to more semantically close (means that distance is high). Like from 0 to 10. In addition, we are showing the most semantically distant distraction concepts just to avoid semantic influence on the process of the filtration.
3. The system contains **x** amount of words which are all presented in referent recognition tasks. The distractors in these tasks are selected from the most distant distractors as possible. Each item is presented two times with two different sets of distractors. The selection is done by choosing first the required amount of nodes from the furthest nodes for two trials, twelve in our case, since each trial requires six distractors and then every other of these is chosen to the trial. The formula for the amount of concepts that we are taking as a distraction ones = N showing times * N distraction concepts.
4. The order of the filtering trials is the same for each learner and familiarity is not considered in choosing the objects. The trials proceed in previously determined order so that none is presented twice in a row. The position of the correct object is randomized. If the learner answers incorrectly to both trials for one word, this word is then considered novel and added to a list of words which would be exposed to in the learning phase. In filtering phase, the learners do not get feedback to their answers.

Fig. 5 shows a typical task in filtering phase where the correct item is ‘Baboon’. As exemplified, the other items are selected from the most distant nodes; none of them are mammals like the baboon.

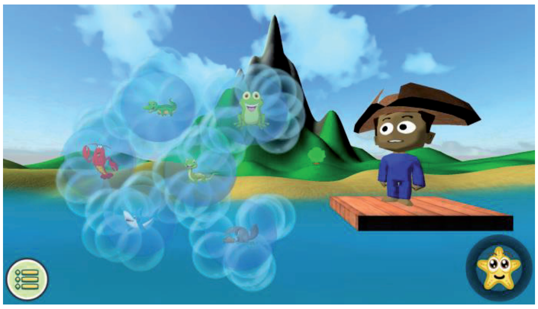


Figure 5. FILTERING PROCESS IN A GAME

D. Learning Process

The learning phase proceeds similarly as the filtering, yet learners are exposed only to the novel words. Before starting the learning phase, the same amount of words is chosen to every learner, so that the amount of trials and exposure time is the same for all players. In practice, this amount is the same as the amount of novel items of the learner that knew most of the words in advance, yet maximum of 20 words to control the total time of the experiment. These novel words are then presented in trials as follows. The experiment group's distractors are chosen from the most similar nodes in the network and the control groups in contrast, from the most distant items. These distractors are chosen in practice similarly as in filtering phase — twelve nodes for each word (six for both trials with distractors), choosing every other for the trials so that the same distractors are not shown with the same item. If familiar items are amongst these words, they are left out and another is chosen instead. Note that it is possible though, that some of the distractors are repeated amongst trials with different target items. Each word is presented first without distractors, immediately followed by two trials with six distractors. If the answer is incorrect, the correct answer is shown and the incorrectness is indicated with a sound and an avatar-animation (frowning, shaking head). If the answer is correct, this is indicated with a sound and an animation with the avatar (smiling, nodding). After all words are gone through, the game proceeds to the testing phase. Fig. 6 shows learning process.

E. Testing Process

In the testing phase, all words, in same order as learning, are presented in the trials with most different or least different words depending on whether the learner is in the control or experiment group. The distractors are then the same ones as in learning phase's first trial. Feedback of correct and incorrect answers is given similarly as in learning phase.

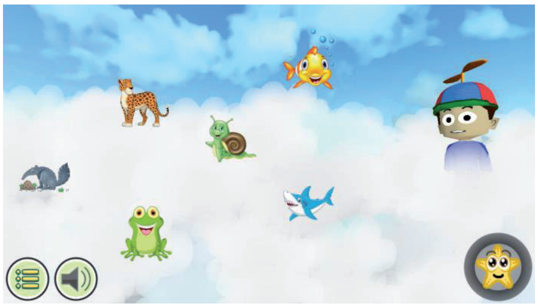


Figure 6. LEARNING PROCESS IN A GAME

VI. CONCLUSIONS AND FUTURE WORK

We presented here a design of a learning game for late talkers, whose prototype is also available. The main ingredient of the proposed system is its adaptivity based on the semantic similarity. The overall game play was divided into filtering, learning, and testing phases.

The goal of this work was a prototype developing of AdapTalk learning game for children for future improvement of their vocabulary. According to our hypothesis that semantic distance between words could influence learning process, we developed learning game, which takes into account those distances.

In the future work, the proposed system will be tested with the children from kindergarten. We plan to compare control and experiment groups of children to study the efficiency of the game design. Especially, we can study the times a child is spending on filtering, learning, and testing phases. Also, the results of the filtration and testing (i.e., correct and incorrect answers) will be analyzed. After analysing the corresponding log files one can make rigor conclusions on the role of semantic similarity in vocabulary learning for young children.

First, testing of AdapTalk should be made with normal children to see if they are recognizing distances. After that to test this learning game with slow learners to see if they have the same learning capabilities as normal children and if semantic distance of the words also influence them. According to the analyzed log files it will be possible to make conclusions and based on them make adaptation process where semantic distance will be regulated according to learners' progress in extending vocabulary.

ACKNOWLEDGMENT

We sincerely thank Professor Heikki Lyytinen for the inspiration, guidance and encouragement in the research work.

REFERENCES

Al-Mubaid, H., Nguyen, H. A., 2006. A cluster-based approach for semantic similarity in the biomedical domain. In 28th Annual International Conference of the IEEE

Engineering in Medicine and Biology New York, USA: IEEE Computer Society.

Alvarez, M. A., Lim, S., 2007. A Graph Modelling of Semantic Similarity between Words, In Proc. 2007 International Conference on Semantic Computing, pp. 355-362.

Beckage, N., Smith, L. and Hills, T., 2010, January. Semantic network connectivity is related to vocabulary growth rate in children. In *Proceedings of the Cognitive Science Society* (Vol. 32, No. 32).

Beckage,Nicole,LindaSmith,andThomasHills.2011.“Small worldsandsemanticnetwork growth in typical and late talkers”. PloS one 6 (5): e19348.

Bin, Shi, et al. Ontology-based measure of semantic similarity between concepts. Software Engineering, 2009. WCSE'09. WRI World Congress on. Vol. 2. IEEE, 2009.

Bouquet, P., Kuper, G., Scoz, M, Zanobini, S., 2004. Asking and answering semantic queries. In Proc. of Meaning Coordination and Negotiation Workshop (MCNW-04) in conjunction with International Semantic Web Conference (ISWC-04), 2004.

Borovsky,Arielle,EricaM.Ellis,JuliaL.Evans,andJeffreyL.Elman.2015. Lexicalleverage: category knowledge boosts real-time novel word recognition in 2-year-olds. Developmental Science. ISSN: 1467-7687. doi:10.1111/desc.12343.http://dx.doi. org/10.1111/desc.12343.

Brusilovsky, P., 1999. Adaptive and intelligent technologies for web-based education. KI 13.4: 19-25.

Brusilovsky, P., Peylo, C., 2003. Adaptive and intelligent web-based educational systems. International Journal of Artificial Intelligence in Education 13.2: 159-172.

Csikszentmihalyi, M., 1990. Flow: The Psychology of Optimal Experience. Harper Perennial, New York.

Euzenat, J., Shvaiko, P., 2013. Ontology matching. Springer-Verlag, Berlin Heidelberg (DE), 2013, isbn 978-3-642-38720-3.

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E. and Hartung, J. P., et al. 1993, The MacArthur Communicative Development Inventories: User’s Guide and Technical Manual (San Diego: Singular)

Gavriushenko, Mariia, Oleksiy Khriyenko, and Iida Porokuokka. Adaptive vocabulary learning environment for late talkers. CSEDU 2016: Proceedings of the 8th International Conference on Computer Supported Education. Vol. 2, ISBN 978-989-758-179-3. SCITEPRESS, 2016.

Gavriushenko, M., Saarela, M., & Kärkkäinen, T. (2017). Supporting Institutional Awareness and Academic Advising using Clustered Study Profiles. In P. Escudeiro, G. Costagliola, S. Zvacek, J. Uhomoibhi, & B. M. McLaren (Eds.), CSEDU 2017 : Proceedings of the 9th International Conference on Computer Supported Education. Vol. 1 (pp. 35-46).

Gifford T., 2013. Take It Personally: Adaptive Learning. retrieved from http://eltjam.com/adaptive-learning/

Hirst, G., & St-Onge, D. (1998). Lexical chains as representations of context for the detection and correction of malapropisms. In C. Fellbaum (Ed.), WordNet: An Electronic Lexical Database (pp. 305–332).

Haase, P., Siebes, R., van Harmelen, F., 2004. Peer Selection in Peer-to- Peer Networks with Semantic Topologies. In Proc. of International Conference on Semantics of a Networked World: Semantics for Grid Databases.

Jiang, J., Conrath, W., 1997. Semantic Similarity Based On Corpus Statistics and Lexical Taxonomy. In Proceedings of International Conference Research on Computational Linguistics.

Kiili, K., 2005. Digital game-based learning: Towards an experiential gaming model. The Internet and higher education 8.1: 13-24.

Kondrak, G., 2005. N-gram similarity and distance. In String processing and information retrieval Springer Berlin Heidelberg: 115-126.

Leacock, C., Chodorow, M., 1998. Combining local context and WordNet similarity for word sense identification, In Fellbaum MIT Press, pp. 265–283, 1998.

Leonard, Laurence B. Children with specific language impairment. MIT press, 2014.

Li, Y., Bandar, Z., McLean, D., 2003. An approach for measuring semantic similarity between words using multiple information sources. Knowledge and Data Engineering, IEEE Transactions on, 15(4), 871-882.

Lyytinen, P, Poikkeus, A-M, Laakso, M-L, Eklund, K, Lyytinen, H., 2001. Language development and symbolic play in children with and without familial risk for dyslexia. J Speech Lang Hear Res 2001; 44: 873–85.

Lyytinen, Heikki, et al. "The development of children at familial risk for dyslexia: birth to early school age." Annals of dyslexia54.2 (2004): 184-220.

Lyytinen, Paula, Kenneth Eklund, and Heikki Lyytinen. Language development and literacy skills in late-talking toddlers with and without familial risk for dyslexia. Annals of dyslexia 55.2 (2005): 166-192.

Lyytinen, Kalle. HCI research: Future directions that matter. AIS Transactions on Human-Computer Interaction 2.2 (2010): 2.

Mokotoff, G., 1997. Soundexing and genealogy retrieved from http://www.avotaynu.com/soundex.htm.

Murray, Meg Coffin, and Jorge Pérez. 2015. Informing and Performing: A Study Comparing Adaptive Learning to Traditional Learning. Informing Science: The International Journal of an Emerging Transdiscipline.

Oxman, S., Wong, W., D. V. X. Innovations, 2014. White Paper: Adaptive Learning Systems. Snapwiz. Integrated Education Solutions, 1.

Preston, J., Frost, S., Mencl, W., Fulbright, R., Landi, N., Grigorenko, E., Jacobsen, L., Pugh, R., 2010. Early and late talkers: school-age language, literacy and neurolinguistic differences. Brain 133 (8): 2185–2195. ISSN: 0006-8950. doi:10.1093/brain/awq163.

Marulis, Loren Marie, and Susan B. Neuman. How vocabulary interventions affect young children at risk: A meta-analytic review. Journal of Research on Educational Effectiveness 6.3 (2013): 223-262.

Marulis, Loren M., and Susan B. Neuman. The effects of vocabulary intervention on young children's word learning: A meta-analysis. Review of educational research 80.3 (2010): 300-335.

Redmond, Sean M., and Mabel L. Rice. "The socioemotional behaviors of children with SLI: Social adaptation or social deviance?." *Journal of Speech, Language, and Hearing Research* 41.3 (1998): 688-700.

Ronimus, Miia and Kujala, Janne and Tolvanen, Asko and Lyytinen, Heikki. "Children's engagement during digital game-based learning of reading: The effects of time, rewards, and challenge", In Computers & Education, Volume 71, 2014, Pages 237-246

Skiada, R., Soroniati, E., Gardeli, A., Zissis, D., 2014. EasyLexia: A mobile application for children with learning difficulties. Procedia Computer Science, 27, 218-228.

Sonwalkar, N., 2005. Adaptive learning technologies: From one-size-fits-all to individualization. EDUCAUSE Center for Applied Research, Research Bulletin(2005).

Simard, P., LeCun, Y., Denker, J. S., 1993. Efficient pattern recognition using a new transformation distance. In Advances in neural information processing systems (pp. 50-58).

Venable, M., 2011. Adaptive Learning Technology: An Introduction.Web-site: http://www.onlinecollege.org/2011/08/30/adaptive-learning-technology-an-introduction/

Wang, L., Zhang, Y., Feng, J., 2005. On the Euclidean distance of images. In Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol.27, no.8, pp.1334-1339, Aug. 2005.

Winkler, W.E., 1999. The state of record linkage and current research problems. In Statistical Research Division, US Census Bureau.

Wu, Z., & Palmer, M. (1994). Verb semantics and lexical selection. In 32nd annual Meeting of the Association for Computational Linguistics (pp. 133–138). Las Cruces, New Mexico: Association for Computational Linguistics.