Virtual Assistants for Learning: A Systematic Literature Review

Regina Gubareva[1] and Rui Pedro Lopes[2][a]
[1]*Taraz State University, Kazakhstan*
[2]*Research Center for Digitalization and Industrial Robotics, Instituto Politécnico de Bragança, Portugal*

Keywords: Virtual Assistant, Learning Management System, Machine Learning, Higher Education, Systematic Literature Review.

Abstract: A problem of students' motivation, engagement and declining interest in the learning process has always existed, contributing to increasing failures and dropouts. This is particularly important among first year students. Freshmen often have difficulties with time management, how to prioritize tasks, and how to study at the university. Because of the increasing number of higher education students, it is impossible to provide individual tutoring and support to every student, to help them manage this first-year indefiniteness and later difficulties. Recent developments in the area of information technology, software engineering, artificial intelligence, machine learning, and big data creates the opportunity for personalized, flexible and adaptable learning environment, accessible anytime, anywhere. One such example is a virtual assistant, a tool that provides assistance to usually boring or repetitive daily activities. In education, a virtual assistant can help organising the study process, manage time, increasing motivation and engagement in the study process. This paper performs a systematic literature review of the use of virtual assistants in higher education. It focuses on the technology that powers them, their features and their impact in the learning process, motivation and productivity, according to the authors.

1 INTRODUCTION

Education is one of the most important aspects of human development, greatly influencing the path of professional development and success (Mesquita et al., 2015; Mesquita et al., 2014). The increasing need for training at the higher-education level, that contributes to the scientific and technological qualification of youth and adults, towards the "creative, innovative and competitive development, with high productivity standards" (Correia and Mesquita, 2006, p. 166) is a huge challenge that higher education institutions face. The admission of students with heterogeneous profiles, with different motivations, academic and social background, with diverse influences create important asymmetries, requiring adequate analysis of their characteristics and profiles to optimize the academic success.

Higher education is, thus, redefining its paradigm, moving from a traditional lectures perspective to an active and emancipating learning process. This requires not only the understanding of the audiences, but also permanent reflection and cooperation between all the involved actors. In fact, although the integration of students in higher education foster their intellectual and ethical behaviour, not many manage to achieve higher levels in this development. This requires considerable effort from institutions to help students to achieve higher intellectual, ethical and professional development.

Adequate pedagogical and training strategies, such as the involvement of teachers in their pedagogical training (Rosado-Pinto, 2008), tutoring, project-based learning (Veiga Simão et al., 2008; Simão et al., 2002) have contributions to the integration and learning of students in higher education (Lopes and Mesquita-Pires, 2013).

To a higher degree, the quality of the learning process depends on student motivation and involvement (Hanus and Fox, 2015), even when facing massification of education, declining completion rates, and, sometimes, incoherence of courses. This, further associated with the students' difficulty in managing study time, contributes to increasing dropout and low academic success. The workload required for a student to achieve the objectives of a particular course is dependent on their work skills, the level of the objectives and contents of each course and the teaching-

[a] https://orcid.org/0000-0002-9170-5078

Gubareva, R. and Lopes, R.
Virtual Assistants for Learning: A Systematic Literature Review.
DOI: 10.5220/0009417600970103
In *Proceedings of the 12th International Conference on Computer Supported Education (CSEDU 2020) - Volume 1*, pages 97-103
ISBN: 978-989-758-417-6

learning methods. The process puts great emphasis on the student, encouraging higher education institutions and academic staff to place students at the centre of their thinking and to help them manage their expectations and be able to consciously and constructively design their learning paths throughout their higher education experience (Lopes et al., 2019; Tenorio et al., 2018).

In this framework, the students' autonomous work assumes a fundamental value in learning. It is important that they know and recognize the objectives, what is being studied, how to define the work tasks and priorities, how to use and to enjoy the numerous information resources, how to write summaries and to prepare reading sheets and reports. In other words, they need to be persistent and manage to take the work to the end. This suggests that students can benefit from aid to help them manage all the complexity associated with the what (what they need to do), when (the amount of time, the deadlines, the schedule organization) and how (the content, exercises and procedures) to learn. This is where a virtual assistant can be useful.

This paper explores, through a systematic literature review, the use and importance of virtual assistant in the motivation and autonomy of higher education students.

2 METHODOLOGY

The main objective of this literature review is to try to understand the role of virtual assistants in higher education, in respect to the impact on the students' time management and autonomy development. Additionally, it is also important to understand the definition of virtual assistants according to the authors and how are they implemented.

This literature review follows the approach suggested by Materla et al. (Materla et al., 2017) and by Subhash and Cudney (Subhash and Cudney, 2018). It is composed of three phases, starting with the planning, followed by the operation (conducting) and dissemination (reporting) phases.

The planning phase included both the definition of the bibliography databases and the selection of the query term. The selected databases were Scopus and Web of Knowledge (WoK), since they provide the most relevant and accurate results. The query used the term `(\virtual assistant" AND \higher education")` for the title, keywords an abstract. The papers were further restricted to the last 20 years, from January 1st, 2000 and December 31st, 2019. Only papers available in the institutional repositories, peer reviewed and written in English were considered.

The second phase (conducting) started with the selection of the relevant papers and exclusion of the remaining. Repeated papers and papers without available full-text PDF were excluded. After these step, a total of 51 papers remained. The process continued with the selection of the most relevant articles, through title and abstract analysis. Non relevant papers were eliminated. Subsequently, the most relevant papers remained, in a total 18 (Table 1).

Table 1: Number of papers during the analysis.

Source	Search results	Repeated and full-text	title and abstract
Scopus	135	32	13
WoK	115	19	4
Total	250	51	17

The third phase consists of the reporting of the results.

3 ANALYSIS AND DISCUSSION

The analysis started with the characterization of the final 17 papers. A content analysis followed, to assess the context and definition of virtual assistants in education, according to the purpose of this work.

In total, papers from 13 countries were found. England was the country with the highest number of papers (3). Russia and Switzerland followed, with 2 papers each. The remaining countries, namely Singapore, Pakistan, Taiwan, Canada, India, France, Bulgaria, Saudi Arabia, Spain and Germany are also mentioned with 1 paper each (Figure 1).

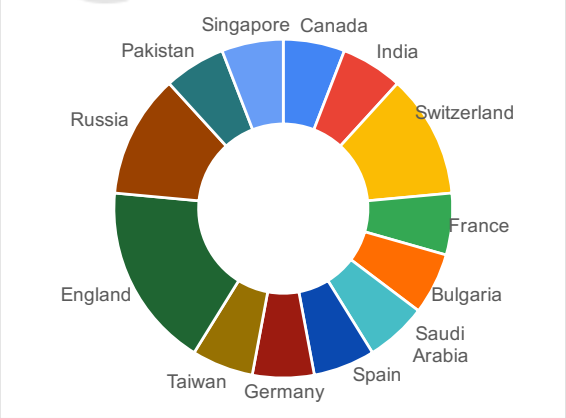

Figure 1: Number of papers per country.

It seems that the number of papers per year has steadily been growing since 2000, with the exception

of 2010 (Figure 2). This indicates that the field is still new and researchers are starting to invest time and effort in the development of this area.

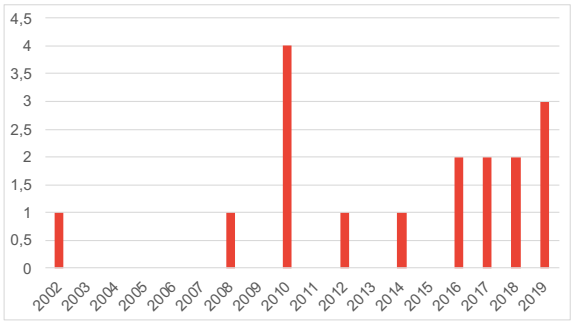

Figure 2: Papers published each year.

Most of the studies originate in the technological area, conducted by scientists from Computer Science, Informatics and Information technologies. Some also fall under the Education and Learning Technologies (Figure 3).

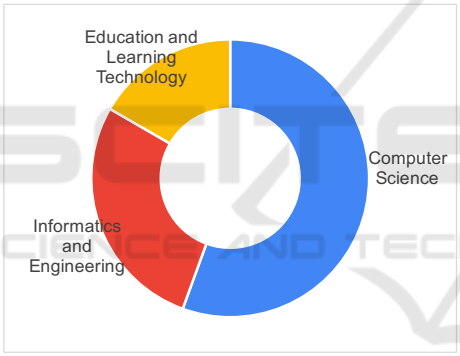

Figure 3: Papers published in each area.

After the characterization, the papers content was analysed with the assistance of text mining algorithms and techniques.

3.1 Text Mining

The PDF files were used to build a text corpus. All the process was performed in R, using the tidytext package (https://github.com/juliasilge/tidytext). The corpus was pre-processed, removing repeated forms, building a dictionary of terms, eliminating irrelevant words and minimizing the number of different words through reduction of inflectional form of the words (stemming).

After this initial step, a histogram of unigrams (terms composed of a single word) and bigrams (terms with two consecutive words) were calculated (Figures 4 and 5).

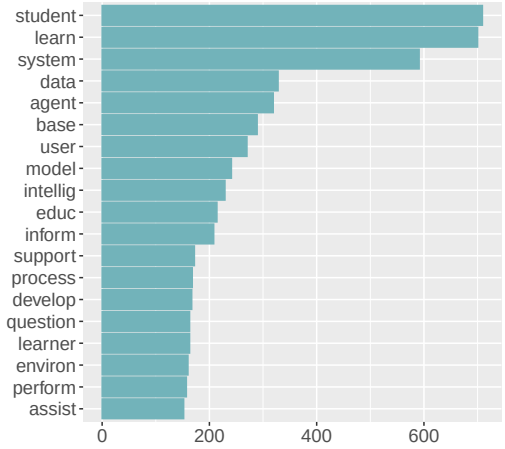

Figure 4: Unigram frequency for the whole corpus.

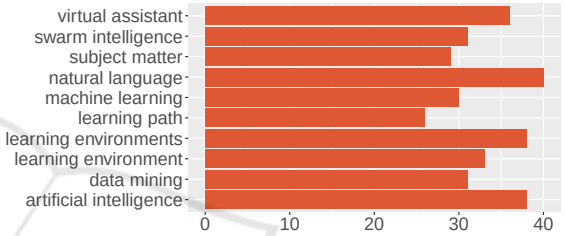

Figure 5: Bigram frequency for the whole corpus.

As expected, both figures suggests that the main focus of the papers is related to technology, with application in education and learning. Most of the words and bi-words refer to student, machine and artificial intelligence, learning environments, virtual assistant, and others. The focus is, clearly, on education, supported by intelligent, virtual assistants, as we expected from the search previously conducted. "Virtual assistant" is the most frequent keyword which was founded in the papers. The second important key phrase is "intelligent learning systems". The intelligent learning system means the learning management system with an embedded virtual assistant, tutor, chatbot or intelligent agent. A really common case when creating a whole system, except for a virtual assistant. Also, frequent concept is a virtual assistant for learning analytics. It means that virtual assistant is an application or web-service which helps students to get their learning analytics. One of the most popular concepts of a virtual assistant is a chatbot. The idea is so admired because there is a desirable opinion that the most convenient way of the student interconnection and learning environment is a voice assistant. There are plenty of studies that create a chatbot. We only take two the most interesting researches (cases). Artificial intelligence is applied more often.

After the global frequency of words and bi-words,

a comparative relevance analysis was made, through the Term Frequency/Inverse Document Frequency (TF-IDF) numerical statistic index. This numerical statistic reflects how important a word is to a document in a collection of papers. In other words, it assesses the most frequent terms and the most relevant terms for each document (Figure 6).

From these figures it is possible to understand what each paper values more, in comparison with the others. This gives a clue about the main focus of each paper, which provides a base for the category identification through a qualitative analysis.

3.2 Content Analysis

The main research question is to understand the influence of virtual assistants on students autonomy and motivation. The resulting papers approach this in several ways. Albalowi and Alhamed (Albalowi and Alhamed, 2017) start by categorizing the most influencing factors for student retention:

1. Academic Integration.
2. Social Integration.
3. Institutional Commitment.
4. Out-of-institution factors.

They conclude that virtual assistants could be applied in the first three factors, thus providing support for students in this issues and helping preventing retention. The first area is academic integration, where learning analytics through machine learning algorithms can help preventing students' difficulties. They use big data and multiple data types to examine students' performance. Two predictive models are created: the first model was built using the traditional structured data to predict the students' final performance, and the second was built using the same data in addition to the students' sentiment scores. The students' sentiment scores were calculated by analyzing their textual feedback using the Stanford Core NLP Natural Language Processing Toolkit. They can support students by providing them learning analytic and failure prediction.

This is further supported by Ciolacu et al. (Ciolacu et al., 2018). The authors use Learning Analytics (LA) for an Early Recognition System with Machine Learning for a personalized email for prompting the students of risk. They applied machine learning algorithms such as Support Vector Machine (SVM) and Neural Networks (NN) to reduce the failure rate in examinations. To motivate, especially first year students, activity diagram of former students are shown. This underlines the activities in the mathematics course during the semester and grades/scores (high scores of those who pass the exam).

The second area is social integration. Lamontagne et al. (Lamontagne et al., 2014) affirm that it is one of the most substantial part of students' success. Coincidentally, it is the most difficult problem to solve. There is no easy way to create chatbots or computer algorithm that can replace real people communication, simulate relationship in group, discuss and actually foster collaboration between students. Sometimes, social integration plays an integral role of the successful education. They even define "a virtual assistant as mainly an information service".

The third area is an institutional commitment. There are a lot of papers which consider especially this area: chatbot, intelligent agents, tutors, and others. The AICMS with Dialog Flow is a system that runs as a messenger end point in Facebook, which takes the text and voice as input and provides answers as text and voice as well. It provides all required information about the college, student information, examination section information, placement cell information, and cultural events (Arun et al., 2019).

Laeeq and Memon (Laeeq and Memon, 2019) presented their "Scavenge" parencite. The Scavenge is an AI-based Intelligent Multi-Agent System (MAS), developed using JADE (Java Agent Development Framework). The authors proposed their variant of the system as three agents platform. The concept was proposed earlier by Zhang et al. (Zhang et al., 2010), where the agents assume the roles of the Learner, the Teacher and the Platform, actively participating in the educational process (Zhang et al., 2010).

Zakharova (Zakharova, 2018) has created a professional development forecast model. She used machine learning algorithms to analyze students' papers. Pre-processing and statistical analysis of texts were carried out to highlight features characterizing the general vocabulary and the use of general scientific and professional terminology. These approaches allows students to get immediate feedback, contributing to better learning and academic success. In fact, this literature review further confirms that the effectiveness of a virtual assistant depends on the quality of feedback.

There are three kinds of feedback elements given by the virtual assistant: evaluation, disturbances and hints. Disturbances are questions that are intended to test the confidence of the students about their work. Hints are documents (text, graphs, formulas, etc.) related to a particular question of the experimental protocol that help students find by themselves possible errors or problems (Geoffroy et al., 2002).

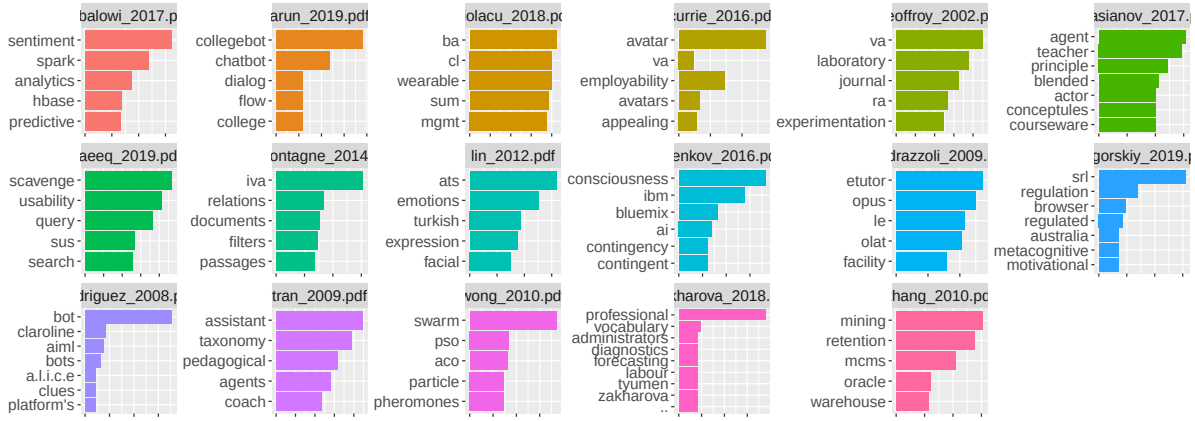

Figure 6: Unigram TF-IDF index.

Pedrazzoli and dall'Acqua (Pedrazzoli and dall'Acqua, 2009) proposed a concept of a platform that recognizes three general levels of advice: a) a reminder of the current target; b) a general description of how to achieve the solution; c) a description of exactly which problem solving action steps should be taken. Each of these three levels may be represented by multiple assistance steps. They assume that a personal assistant should be able to: recognize a large variety of student solutions; diagnose student "Subject Matter" understanding and recommend target oriented, optimized "learning approach adaptations"; tailor tutorial actions accordingly; support collaboration; support specific forms of adaptation for collaboration activities, like recommending suitable collaborators and actions; adapt the interface to facilitate collaboration activities (enforce specific roles and rules); advise students how to interact efficiently; reasoning, specify techniques to acquire and propose additional knowledge material about a domain or subject matter; use the knowledge base to solve problems in that domain or subject matter; support educational workflow sequences.

In addition to supporting the workflow of information, virtual assistants can also be used as tools for time management and motivation. Each student has a specific profile, with different academic, social and economic backgrounds. Each student will have his own time, effort and process for learning. A teaching-learning process that is too much different and incompatible with the student's profile can lead to dropout. Thus, dropout prevention can be approached with personalized learning environments, to best fit each student's abilities and restrictions. Tran et al. (Tran et al., 2009) suggests a personal tutor, an intelligent assistant that follows students during their learning and helps students to overcome their difficulties. If necessary, it can refer the student to other support mechanisms, such as counsellors, placement officers, examination officers, accommodation officers etc., each of whom can assume the support role in their specific context.

There is a dependency between the implementation of virtual assistant and student learning motivation. Surprisingly, students had almost no interest in using contact via phone or social media. According to the survey, students currently prefer communicating with tutors face-to-face and by email (Currie et al., 2016). According to these authors, the virtual assistant should be implemented as a voice service, 3D agent or avatar, to foster motivation. They are more user-friendly, handy and generally more interesting for students. The results of the study indicated that "students expressed positive interest in avatars and their motivation to learn".

A classical representation of virtual assistant consists the roles of teacher and encourager. In one step further, establishing the three-agent learning as the learner-system-teacher, the system might be more than just a "mentor". According to the study experience of using three part system shows that the engagement and the involvement of the students that the immediate proper feedback and problem-based learning provide give measurable and profound advantage over the ordinary education process (Khasianov et al., 2017).

Lin et al. (Lin et al., 2012) focus on the understanding of student's states. The system is based on dual-mode operation: facial expression recognition, and text semantics as the main elements in affective computing to understand users' emotions. Text semantics are used to understand learner's learning status, and the results contribute to course management agents in order to choose the most appropriate teaching strategies and feedback to the users. According to the authors, emotions is an important factor for learn-

ers' motivation, whereas motivation plays a vital role in knowledge development.The use of emotion recognition in the learning system had good feedback from students. According to the study there was an increase in user interest in learning digital art.

Nenkov et al. (Nenkov et al., 2016) used artificial intelligent technologies to simplify scheme of academic interaction. The idea is to use special AI agent as third part in Facebook messenger. The interaction between the teacher and the student is carried out in the online environment (in social net, for example) through mediation of an AI agent (chatbot) which acts on the basis of knowledge of the LMS services. This AI agent is likely to show competencies in a narrow area (which is limited by the content of the course topic) in combination with the general communication and intelligence skills. The feature enables for the students to integrate learning in familiar social networks and for the teachers to reduce the burden and relieve them from monotonous work.

Rodriguez et al. (Rodriguez et al., 2008) use AI technologies in a couple of chatbot: T-BOT and Q-BOT. These are two virtual assistants designed to tutor and evaluate students into an e-learning platform. Both have been designed as PHP modules so they can be easily integrated into e-learning platforms like Moodle. The Q-BOT and T-BOT entities gives students opportunity to evaluate and monitor their progress as a real one, guiding and tutoring students in their access to platform's resources, T-BOT is able to answer students' questions about different subjects using natural language. They are both great instruments integrated into e-learning system, designed to create friendly and clear learning environment. It should be noted that this is an essential requirement for such systems, where the main factor influencing student motivation is convenience and comprehensibility.

Another requirement is an adaptive and personalised assistant. Adaptive navigation helps students locate and navigate in information hyperspace. Adaptive presentation adapts the contents or display of a page according to the user's profile. Adaptive collaboration helps learners to find the most suitable helpers or collaborators (Wong and Looi, 2010). To meet these requirements the authors used swarm intelligence, a set of computing algorithms in the form of multi-agent systems that simulate how swarms of insects or birds move or work. The key feature of swarm intelligence is "to balance exploitation (of the known best solution) and exploration (of the unknown solutions)". This feature in such swarm systems facilitates adaptivity in dynamic or unstable environment. According to the authors, swarm intelligence has proven efficient in solving highly complex optimisation problems or pattern recognition and plays a key role to improve adaptivity of educational systems.

Pogorskiy et al. (Pogorskiy et al., 2019) discuss the design of a virtual learning assistant which consists of an extension to the Chrome web browser to be used as an instrument to support self-regulated learning. The extension allows for the collection of data on its users' web sessions, and interacts with the user through pop-up notifications and the extension's dashboard located in the Chrome menu. Micro-randomized notification can be useful for self-regulation.

4 CONCLUSIONS

In this paper we have explored the research in the field of virtual assistants in higher education. The review provides an array of uses and techniques for developing virtual assistants and how they can be used for education support.

Student support can be categorized depending on the role of virtual assistant in four types: digital tutor, the digital secretary, the motivator agent, the mentor agent. The tutor's role is to help students through the learning process The digital secretary helps organising the learning process, solving administrative issues, and remind deadlines. The motivator helps integrating in social environments, to help overcome stress and anxiety, remind the main aim and motivate students. Finally, the mentor gives a general description of how to achieve the solution and description of exactly which problem solving action steps should be taken.

Virtual assistants are becoming popular and useful technology, with a variety of advantages, contributing to automation of tasks and providing support for students in time-management, access to information and communication facilitation. The technology is still in its infancy. There are many aspects that are necessary to improve to make virtual assistant effective in student motivation and engagement.

REFERENCES

Albalowi, A. and Alhamed, A. (2017). Big data and learning analytics in higher education: Demystifying variety, acquisition, storage, NLP and analytics.

Arun, K., Nagesh, A., and Ganga, P. (2019). A Multi-Model and Ai-Based Collegebot Management System (Aicms) for Professional Engineering Colleges. *International Journal of Innovative Technology and Exploring Engineering*, 8(9):2910–2914.

Ciolacu, M., Tehrani, A. F., Binder, L., and Svasta, P. M. (2018). Education 4.0 - Artificial Intelligence Assisted Higher Education: Early recognition System with Machine Learning to support Students' Success. In *2018 IEEE 24th International Symposium for Design and Technology in Electronic Packaging (SIITME)*, pages 23–30, Iasi. IEEE.

Correia, A. M. R. and Mesquita, A. (2006). *Novos públicos no ensino superior desafios da sociedade do conhecimento*. Edições Sílabo, Lisboa. OCLC: 494458369.

Currie, E., Harvey, P. H., Daryanani, P., Augusto, J. C., Arif, R., and Ali, A. (2016). An investigation into the efficacy of avatar-based systems for student advice. *EAI Endorsed Transactions on e-Learning*, 3(11):151450.

Geoffroy, F., Aimeur, E., and Gillet, D. (2002). *A virtual assistant for web-based training in engineering education*, volume 2363 of *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*.

Hanus, M. D. and Fox, J. (2015). Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance. *Computers & Education*, 80:152–161.

Khasianov, A., Suleymanov, D., and Marchenko, A. (2017). Three agent platform approach for digital education environment. pages 8580–8587, Valencia, Spain.

Laeeq, K. and Memon, Z. (2019). Scavenge: an intelligent multi-agent based voice-enabled virtual assistant for LMS. *Interactive Learning Environments*.

Lamontagne, L., Laviolette, F., Khoury, R., and Bergeron-Guyard, A. (2014). A framework for building adaptive intelligent virtual assistants. *Proceedings of the IASTED International Conference on Artificial Intelligence and Applications, AIA 2014*, pages 342–349.

Lin, H.-C. K., Wang, C.-H., Chao, C.-J., and Chien, M.-K. (2012). Employing Textual and Facial Emotion Recognition to Design an Affective Tutoring System. *Turkish Online Journal of Educational Technology*, 11(4):418–426. WOS:000309531600042.

Lopes, R. P., Mesquita, C., de Góis, L. A., and dos Santos Júnior, G. (2019). Students' Learning Autonomy: A Systematic Literature Review. pages 5958–5964, Palma, Spain.

Lopes, R. P. and Mesquita-Pires, C. (2013). Novos públicos no Ensino Superior Politécnico: desafios de formação. In Membiela, P., Casado, N., and Cebreiros, M. I., editors, *III International Conference on Higher Education Teaching*, pages 535–539. Educación Editora, Ourense, Spain.

Materla, T., Cudney, E. A., and Antony, J. (2017). The application of Kano model in the healthcare industry: a systematic literature review. *Total Quality Management & Business Excellence*, 0(0):1–22.

Mesquita, C., Lopes, R. P., García, J. l., and de la Cruz del Río Rama, M. (2015). First Contact with the Word of Work: The Competence Built in the Teaching Practices. In Peris-Ortiz, M. and Merigó Lindahl, J. M., editors, *Sustainable Learning in Higher Education SE - 6*, Innovation, Technology, and Knowledge Management, pages 75–87. Springer International Publishing.

Mesquita, C., Lopes, R. P., García, J. l., and Rama, M. d. l. C. d. R. (2014). Pedagogical Innovation in Higher Education: Teachers' Perceptions. In Peris-Ortiz, M., Garrigós-Simón, F. J., and Gil Pechuán, I., editors, *Innovation and Teaching Technologies: New Directions in Research, Practice and Policy*, pages 51–60. Springer International Publishing, Cham.

Nenkov, N., Dimitrov, G., Dyachenko, Y., and Koeva, K. (2016). Artificial intelligence technologies for personnel learning management systems. pages 189–194.

Pedrazzoli, A. and dall'Acqua, L. (2009). OPUS One An Artificial Intelligence-Multi Agent based Intelligent Adaptive Learning Environment (IALE).

Pogorskiy, E., Beckmann, J., Joksimovic, S., Kovanovic, V., and West, R. (2019). Utilising a Virtual Learning Assistant as a Measurement and Intervention Tool for Self-Regulation in Learning. pages 846–849.

Rodriguez, E., Burguillo, J., Rodriguez, D., Mikic Fonte, F., González Moreno, J. C., and Novegil, J. (2008). Developing virtual teaching assistants for open e-learning platforms. pages 193–198.

Rosado-Pinto, P. (2008). Formação pedagógica no ensino superior.: O caso dos docentes médicos. *Revista de Ciências da Educação*, (7):111–124.

Simão, J. V., Costa, A. d. A., and Santos, S. M. d. (2002). *Ensino superior: uma visão para a próxima década*. Gradiva, Lisboa. OCLC: 917200006.

Subhash, S. and Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior*, 87:192–206.

Tenorio, M., Reinaldo, F., Esperandim, R., Lopes, R., Gois, L., and Dos Santos Junior, G. (2018). Céos: A collaborative web-based application for improving teaching-learning strategies. In *Advances in Intelligent Systems and Computing*, volume 725, pages 107–114.

Tran, T. D., Bouchon-Meunier, B., Marsala, C., and Putois, G.-M. (2009). An Intelligent Assistant to Support Students and to Prevent them from Dropout. In *CSEDU*.

Veiga Simão, A. M., Flores, M. A., Fernandes, S., and Figueira, C. (2008). Tutoria no ensino superior: concepções e práticas. *Sísifo/Revista de Ciências da Educação*, (7):15.

Wong, L.-H. and Looi, C.-K. (2010). Swarm intelligence: New techniques for adaptive systems to provide learning support. *Interactive Learning Environments*, 20.

Zakharova, I. (2018). Machine Learning Methods of Providing Informational Management Support for Students' Professional Development. *The Education and science journal*, 20:91–114.

Zhang, Y., Oussena, S., Clark, T., and Kim, H. (2010). Use Data Mining to Improve Student Retention in Higher Education - A Case Study. In *Proceedings of the 12th International Conference on Enterprise Information Systems*, pages 190–197, Funchal, Madeira, Portugal. SciTePress - Science and and Technology Publications.