31 Identifying Risks Factors of Students' Failure in e-Learning Courses

Gewrgakopoulos
Piraeus University of Applied Sciences
igteipir1979@gmail.com

C. Kytagias
Piraeus University of Applied Sciences
ckyt@puas.gr

Y. Psaromiligkos
Piraeus University of Applied Sciences
yannis.psaromiligkos@puas.gr

M. Chalikias
Piraeus University of Applied Sciences
mchalikias@hotmail.com

Dr. K. Artikis
Piraeus University of Applied Sciences
ctartikis@gmail.com

31.1 Abstract

The radical changes that 21th century has brought about in the territory of education have increased the requirements of the underlying stakeholders for smarter tools capable to provide continuous and sophisticated feedback of the educational process.
Decision Makers in this area need to make decisions at various points as well as at multiple levels during the life cycle of the educational system. Learning Management Systems (LMS) can provide decision support services which can be used to increase the learning effectiveness of the new mode of learning as well as the efficient organization of the institutional resources. Identifying students at risk is a major problem. In this paper we provide a framework and detailed case studies for identifying risks factors of students' failure in e-Learning courses and a proposal of how an LMS can be transformed into a Warning System and provide decision support service to Decision Makers.

31.2 Key words:
e-learning courses, risk identification, risk factors, learning management systems.

31.3 Introduction

Identifying students at risk is a major problem in e-learning courses and there are already various approaches and methods in the literature that analyze various issues such as issues

from a sociological standpoint (see Seidman (2005) and Tinto (2012)) or issues from academic performance, demographics, and engagement (see DeBerard et al. (2004), Zhang et. al. (2004), (Aguiar et. al., 2014)). In this paper we provide a framework and case studies with promising results in order to identify risks factors of students' failure in e-Learning courses and a proposal of how an LMS can be transformed into a Warning System and provide decision support services to Decision Makers. Considering the large amount of e-learning courses delivered today in various forms (e.g. MOOCs) we believe that there is great necessity for decision makers to conduct a risk analysis process in order to identify the risk factors of students' failure in these courses and to enhance the effectiveness in this de facto trend in e-learning.
Our framework is described in section 2. Section 3 describes analytically the application of our approach in two real case studies at Piraeus University of Applied Sciences. Finally, conclusions and future research directions are provided in the last section.

31.4 Identifying Risks factors in e-learning courses

31.4.1 Our methodology

Our methodology combines and adapts steps from the risk management process. As it is depicted by the below figure, we propose four phases:

1. Preparation Phase
2. Risk Analysis
3. Warning System Generation
4. Risk Control

The first phase constitutes a preparatory phase before the risk analysis. At this phase decision maker defines the problem. Risk analyst, in cooperation with decision maker models the risk problem, defining risk, the area of risk, the adverse effects of risk, the target group which is affected by the risk and the risk management options, creating a plan of how the specific risk could be controlled. After the risk definition, the decision maker along with the risk analyst review the available data and record new data needed for the management of the risk problem from the learning environment (for example from the log files of a Learning Management System). In the Risk Analysis phase, the risk analyst uses the preparation phase outcome to come up with a model for risks factors identification and prioritization. In the Warning System Generation phase, the risk analyst uses the risks factors identification model to generate a model for prediction purposes. The prediction model determines the requirements for the generation of the warning system. Finally, in the Risk Control phase, the risk analyst delivers the warning system product to the decision maker who puts the warning system into action in order to validate it. After that, risk analyst in cooperation with decision maker check whether risk is controlled through the utilization of the warning system. On the occasion that risk is not controlled, the entire process will be reviewed (see figure 2). In any other occasion, decision maker has managed to control the underlying risk and can provide reports of the improvement of the learning process.

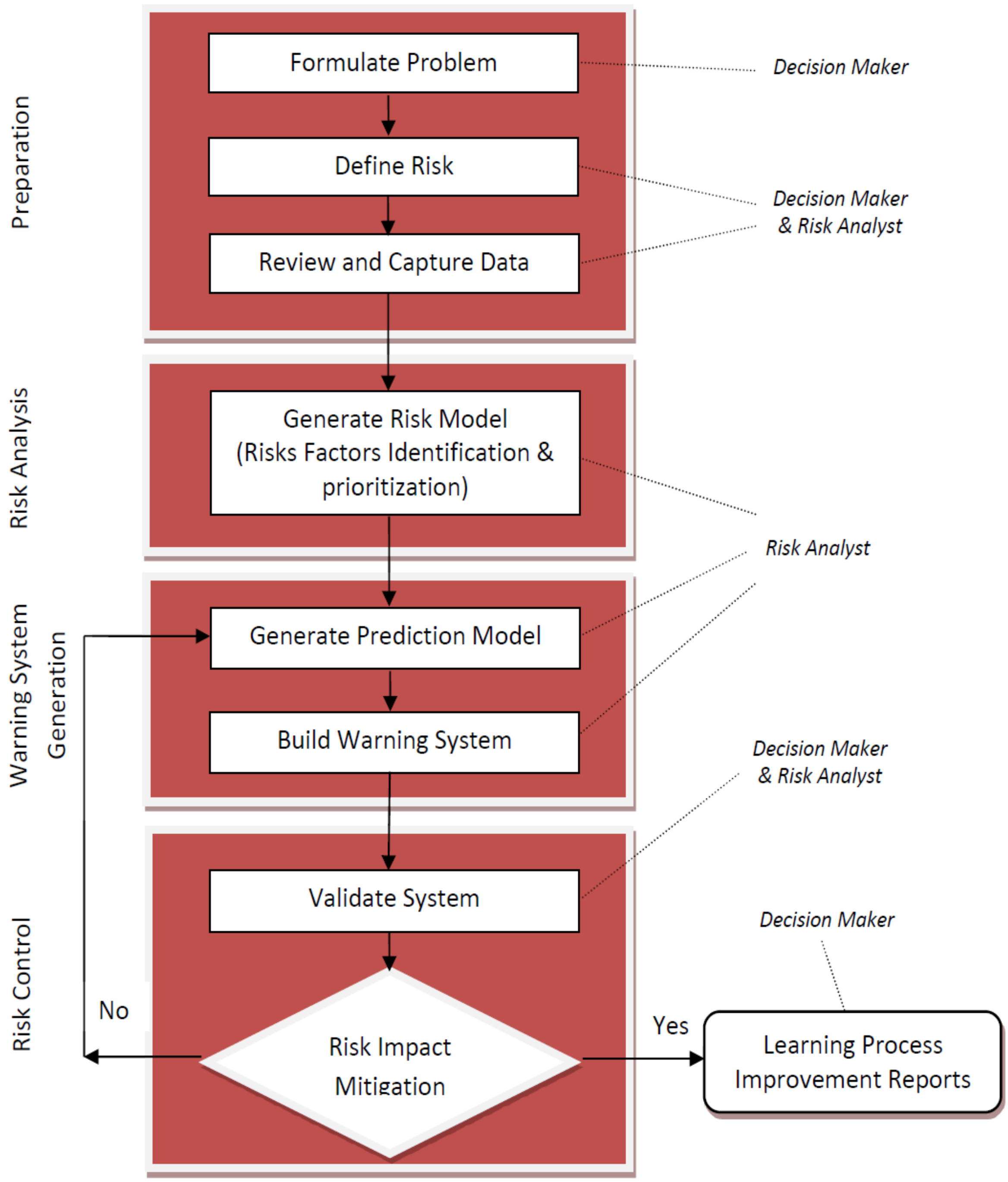


Figure 2. Our methodology

31.5 <u>Applying our methodology to e-learning courses</u>

31.5.1.1 *Case-study 1 Description*

In this section we describe an application of our methodology through a case study that concerns a specific e-learning course offered by Business Administration Department at Piraeus University of Applied Sciences. The course consists of a set of activities distributed in nine (9) sections. The activity types included in each section of the course were: Interactive multimedia learning material developed by the use of the authoring tool "Articulate Storyline 2" (Articulate, 2016), being implemented as a SCORM activity with the SCORM reporting options of "completed", "incomplete" and "not attempted", Video recorded lecture material which had been uploaded to YouTube but it was packaged in Moodle as a single SCORM learning object in order to have the previous mentioned reporting options as well as the total time spent each student in the activity and Self-assessment questions developed by the use of the authoring tool "Articulate Storyline 2" (Articulate, 2016). The underlying activity used the grading method of “High Score”. A student was granted with 3 trials and the final grade was determined by the maximum grade of the trials in a percentile form. Additionally, the students were provided with feedback and support through a forum which was designed to give students the opportunity to pose questions that could be answered not only by the educator but also by their colleagues. A student was considered to have been adequately participated into the forum only on the ground that he/she had made at least two posts. Thereby, in order to complete the course, a student was expected to study the interactive multimedia material included in the nine sections, watch the videos with the recorded lectures, attempt to answer the self-assessment questions and participate into the forum. The students’ performance was controlled through a final online test. Thereby, students could culminate the course successfully only in the case that they had achieved a final grade (in the final online test) greater or equal to 5.

31.5.1.2 Case Study 2 Description

The second case study was referred to an e-learning course having the same structure with the e-learning course described in the first case study, having also being designed in the same way. Nevertheless, it is important to stress on the fact that in the second e-learning course interactive material, self-assessment exercises and videos were divided into 12 sections, instead of 9 (see case study 1). Moreover, the number of students that had been enrolled into the second e-learning course was 234, slightly greater in comparison to the corresponding number of the first e-learning course. The methodology was applied to both e-learning courses in the same way, since it’s a generic methodology, applicable to any e-learning course. The results of both case studies are presented into section 3.3.

31.5.1.3 *Methodology Application*

During the preparation phase of the proposed risk management methodology the decision maker needs to formulate the problem and identify the risks. In our case, the educational problem is the improvement of the underlying course through the improvement of students’ performance and more specifically through the control of students' failure. Thereby, in our case, the risk is the students’ failure. The risk factors identification area is the students’ engagement on the ground that students' engagement has proved to positively correlate to performance. In a more elaborate detail our research has been focused on the students' behavioral engagement denoting the way students behave in terms of an e-learning course (Fredricks, 2004), owing to the fact that an LMS can provide us with such meaningful data in the territory of students' behavioral engagement. These data are shown in the table 1. Appropriate reports of the above data were produced in order to perform some visualization analysis before proceeding to the risk analysis phase. In this case, the technique for the reports generation was based on a special plugin we have developed under the Moodle system (Kytagias et al., 2015).

Table 1: Data captured

Percentage of interactive material parts studied (completed)
Percentage of self-assessment exercises parts completed
Percentage of videos parts watched (completed)
Total number of activities completed
Total number of activities started
Total posts in forum (new posts and follow- up posts)
Total discussions viewed in forum
Time spent on interactive material completion
Time spent on self-assessment exercises' completion
Time spent on videos watch (completion)
Time spent on forum
Total time spent on system
Total logins into the system

During the risk analysis phase we decided to use a binary logistic regression method in order to identify and prioritize risk factors. We have modeled the dependent variable "strisk" (Leah P. Macfadyen, Shane Dawson. (2009)) as the variable that describes the students who are about to phase the risk of not completing the course successfully. There are 2 values ordained in regard to two states: The state "0" denotes students who are not about to face the previously cited risk, whereas the state "1" denotes students who are about to face the referred risk. The state "1" holds true in the case where students' final grade is below the numeric threshold of 5. The state "0" holds true in any other occasion. Detailed results of the application of the binary logistic regression method are presented in the next section. We used the outcome of the risk factors identification process to generate a model to predict students' bad performance. In our case, this step was completed by carrying out a discriminant analysis where we came up with the students' classification into two groups: students who are about to fail the course and students who are about to pass the course. The scores of the discriminant functions were used during the course period to predict whether a student is about to face the risk of not completing the course successfully. That verification process had been completed for the prediction models in terms of both case studies. The warning system could be generated after picking the most suitable prediction model out of the alternative prediction models of both case studies. The verification of the alternative prediction models and the final selection of the prediction model are presented into section 3.4. The outcome from the discriminant analysis could be used to generate a "warning system". That could be achieved only on the ground that the prediction model (see discriminant analysis) having been verified. The verification of the alternate prediction models for both case studies is presented into section 3.4. We believe that learning management systems such as Moodle LMS should provide such a service. In our case, a special plugin is needed that will perform the necessary calculations according to the prediction model for each student and generate suitable messages to students at risks (London et al., 1999; Smith & Ragan, 2005). We are at the process to integrate these calculations to our new plugin we mentioned before for two reasons: (a) the underlying plugin already captures all relevant students' data presented in Table 1 and (b) a plugin inside Moodle fits very well to the purpose of an LMS and provides an easy, one-stop maintenance and reusability of resources. The objective of a warning system generation is the control of the risk.

31.6 Results

31.6.1 Case Study 1 Outcome

Binary logistics regression

The variables determined in Table 1 were used to carry out a binary logistics regression in order to come up with a model for the identification of risks factors. The variables of Table 1 were candidates to constitute risks factors. The regression model will decide which of them should be deemed as risks factors. Table 2 shows the variables participating into the model significantly.

	B	Sig	Exp(B)
Percentage of interactive material parts studied (completed)	-19.974	0.000	0.000
Percentage of self-assessment exercises parts completed	-4.608	0.012	0.010

Table 2: Variables participating into the model significantly

Column Sig, explains that the variables participating into the model significantly are: percentage of interactive material parts studied and percentage of self-assessment exercises parts completed. The generated model enables us not only to identify the risks factors but also to come up with risk factors prioritization. That process is being carried out by calculating the contribution of risks factors to the risk. That contribution is being calculated through the contribution of risks factors to the probability of risk occurrence. Column B of table 2 indicates that one unit increase in the percentage of interactive material parts studied leads to 19.974 units decrease in the logarithm of probabilities. In parallel manner, one unit increase in the percentage of parts of self-assessment exercises parts completed leads to 4.608 units decrease in the logarithm of probabilities. Hence, through the risks factors contribution to the probability of risk occurrence, the cardinal risk factor is the insufficient study of multimedia material (not all sections of the SCORM multimedia material studied). Another risk factor but with slighter contribution to risk occurrence is the insufficient completion of self-assessment exercises (not all sections of self-assessment exercises completed).

Discriminant Analysis

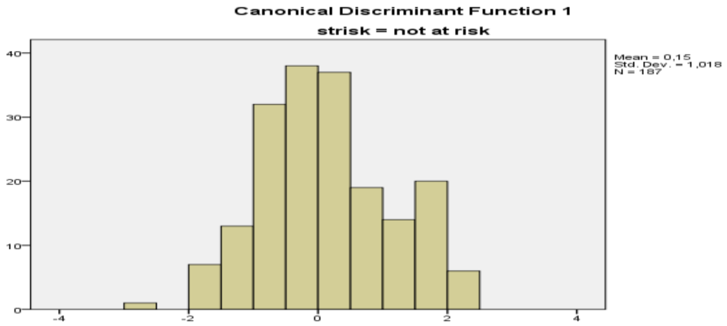
After a discriminant analysis having been conducted, a prediction model of students' critical performance has been generated leading into the generation of two functions, one for students who might not face the underlying risk and one for students who might face the underlying risk. Table 3 gives the discriminant functions coefficients.

Classification Function Coefficients

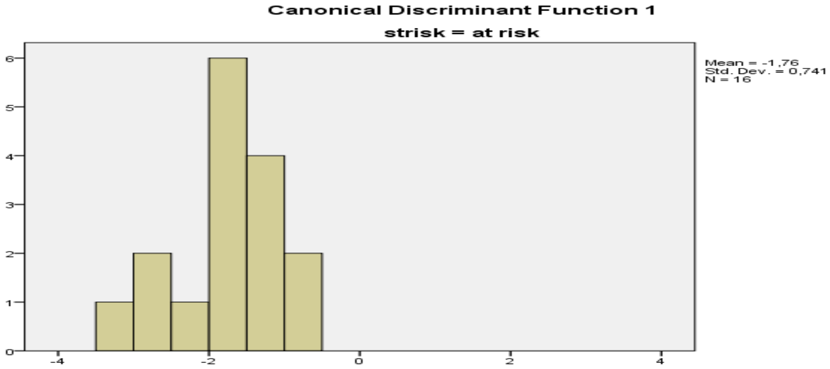
	strisk	
	not at risk	at risk
Percentage of interactive material parts studied (completed)	30,331	19,588
Percentage of self-assessment exercises parts completed	21,187	16,403
(Constant)	-19,871	-10,194

Table 3: Discriminant Functions coefficients

The values of the discriminant functions are well indicated into the below graphs:



Graph 1: Classification function for students not being at risk



Graph 2: Classification function for students being at risk

Classification Results[a]

		strisk	Predicted Group Membership		Total
			not at risk	at risk	
Original	Count	not at risk	161	26	187
		at risk	1	15	16
	%	not at risk	86,1	13,9	100,0
		at risk	6,3	93,8	100,0

a. 86,7% of original grouped cases correctly classified.

Table 4: Correct classification percentage

It is important to highlight that as it is shown into the table 4, the correct classification percentage reaches the amount of 86.7, denoting that discriminant functions achieve great classification.

31.6.2 Case Study 2 Outcome

Binary logistics regression

The variables determined in Table 1 were used to carry out a binary logistics regression in order to come up with a model for the identification of risks factors. The variables of Table 1 were candidates to constitute risks factors. The regression model will decide which of them should be deemed as risks factors. Table 5 shows the variables participating into the model significantly.

	B	Sig	Exp(B)
Percentage of interactive material parts studied (completed)	-10.177	0.000	0.000
Percentage of self-assessments parts completed	-4.759	0.007	0.009

Table 5: Variables participating into the model significantly

Column Sig, explains that the variables participating into the model significantly are: percentage of interactive material parts studied and percentage of self-assessment exercises parts completed. The generated model enables us not only to identify the risks factors but also to come up with risk factors prioritization. That process is being carried out by calculating the contribution of risks factors to the risk. That contribution is being calculated through the contribution of risks factors to the probability of risk occurrence. Column B of table 5 indicates that one unit increase in the percentage of interactive material parts studied leads to 10.177 units decrease in the logarithm of probabilities. In parallel manner, one unit increase in the percentage of parts of self-assessment exercises parts completed leads to 4.759 units decrease in the logarithm of probabilities. Hence, through the risks factors contribution to the probability of risk occurrence, the cardinal risk factor is the insufficient study of multimedia material (not all sections of the SCORM multimedia material studied). Another risk factor but with slighter contribution to risk occurrence is the insufficient completion of self-assessment exercises (not all sections of self-assessment exercises completed)

Discriminant Analysis

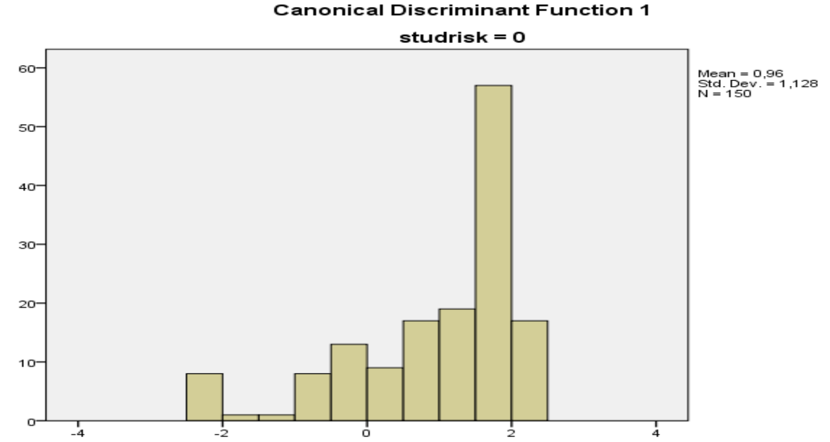
After a discriminant analysis having been conducted, a prediction model of students' critical performance has been generated leading into the generation of two functions, one for students who might not face the underlying risk and one for students who might face the underlying risk. Table 6 gives the discriminant functions coefficients.

Classification Function Coefficients

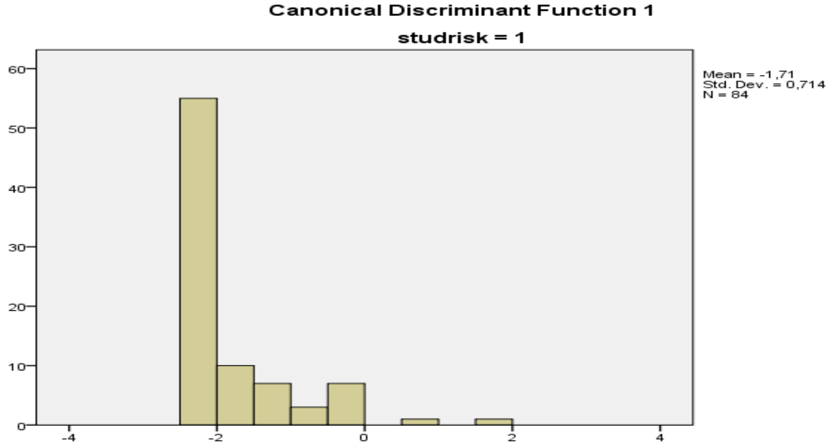
	studrisk	
	,00	1,00
Percentage of interactive material parts studied	11.803	-0,233
Percentage of self-assessment exercises parts completed	13.385	3,381
(Constant)	-5,343	-0,788

Table 6: Discriminant Functions coefficients

The values of the discriminant functions are well indicated into the below graphs:



Graph 3: Classification function for students not being at risk



Graph 4: Classification function for students being at risk

Classification Results[a]

		studrisk	Predicted Group Membership		Total
			,00	1,00	
Original	Count	,00	126	24	150
		1,00	5	79	84
	%	,00	84,0	16,0	100,0
		1,00	6,0	94,0	100,0

a. 87,6% of original grouped cases correctly classified.

Table 7: Correct classification percentage

It is important to highlight that as it is shown into the table 7, the correct classification percentage reaches the amount of 87.6, denoting that discriminant functions achieve great classification.

31.6.3 Towards a Warning System

31.6.3.1 *Selecting the Suitable Prediction Model*

The alternative prediction models of both case studies were verified in terms of their correct prediction percentage. In a more elaborate detail, the scores of the discriminant functions of both prediction models were calculated during the period of another specific common course, before the final examination. The classification results of the discriminant functions were compared to the classification of students after their final examination's grade. The first prediction model achieved a 74% correct classification percentage whereas the second prediction model achieved a 92 % correct classification percentage. Thereby, the second prediction model should be selected to constitute the base of a warning system generation.

31.7 Conclusions

The outcome of the binary logistics regression for both case studies has proved that the measurable risks factors for courses with that specific design, having contribution to risk occurrence are: insufficient study of interactive material and insufficient completion of self-assessment exercises. It is important to stress on the fact that factors associating with time data was not proved to be decisive factors that could affect students' performance critically on the ground that variables related to time had insignificant participation into the regression model. Similar results have been reported in the literature. In this paper we described a risk management framework in order to identify risks factors of students' failure in e-Learning courses. We applied this framework on two case studies where the data captured by an LMS (Moodle) gave promising results to the identification process of students at risk. Our implementation framework gave also promising results of how an LMS could be transformed into a Warning System in order to provide decision support services to Decision Makers. The underlying framework can be applied to any e-learning course for identifying and prioritizing the risk factors of the students' failure. Decision Makers can use the underlying framework not only to identify risk factors of students' failure in their courses but also to redesign their courses by analyzing the above factors in other dimensions such as the dimension of students' learning preferences.

31.8 References

Aguiar E., Chawla N. V., Brockman J., Ambrose G. Alex, Goodrich V., (2014). Engagement vs performance: using electronic portfolios to predict first semester engineering student retention. Proceedings of the Fourth International Conference on Learning Analytics And Knowledge, Pages 103-112, Indianapolis, Indiana, USA — March 24 - 28, 2014.

Articulate, (2016). Articulate Storyline 2. Available at: https://articulate.com/p/storyline-2.

Campbell, J.P. and Oblinger, D.G. (2007) Academic Analytics, Educause.

Carlson, P.A. (1998). Advanced Educational Technologies – Promise and Puzzlement, Journal of Universal Computer Science 4(3).

David Vose. (2008). Risk Analysis: A Quantitative Guide, 3rd Edition.

DeBerard M. S., Spielmans G., and Julka D., (2004). Predictors of academic achievement and retention among college freshmen: A longitudinal study. College Student Journal, 8(1):66{80, 2004.

Finn, J. D. (1989). Withdrawing from school. Review of Educational Research, 59, 117 – 142.

Ford, P., Goodyear, P., Heseltine, R., Lewis, R., Darby, J., Graves, J., Sartorius, P., Harwood, D., & King, T. (1996). Managing change in higher education: A learning environment architecture. London, UK: Open University Press.

Fredericks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. Review of Educational Research, 74, 59 – 109.

G. Zhang, T. J. Anderson, M. W. Ohland, and B. R. Thorndyke. (2004). Identifying factors inuencing engineering student graduation: A longitudinal and cross-institutional study. Journal of Engineering Education, 93(4):313{320, 2004.

Gagne, R. (1987). Instructional technology foundations. Hillsdale, NJ: Lawrence Erlbaum.
James J. Appleton, Sandra L. Christenson, Michael J. Furlong.(2008). Student Engagement with School: Critical Conceptual and Methodological Issues of the Construct. Psychology in the Schools, Vol. 45(5).

Kytagias C., Liomas Y., Livanis T., Psaromiligkos Y., Dimakos G., (2015). "Mining the Log Files of Moodle: Towards Moodle Analytics", International Conference 'Science in Technology' SCinTE 2015, Athens 5-7 November 2015

Larusson A ., White B., (2014). Learning Analytics From Research to Practice
Leah P. Macfadyen , Shane Dawson. (2009). Mining LMS data to develop an ''early warning system" for educators: A proof of concept.

Lindner, R. 2001. Expertise and Role Identification for Learning Environments (ERILE), Proposed Standard Draft for German DIN NI-36, available at: http://www.igd.fhg.de/~lindner/PROMETEUS/SIG-DESIGN_Meeting - Point.html

London, M., Larsen, H. H., & Thisted, L. N. (1999). Relationships between feedback and self-development.Group and Organizational Management, 24(1), 5–27.

Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. American Educational Research Journal, 37, 153 – 184.

Molak, V. (1997).Fundamentals of risk analysis and risk management.

Moore, M. G., and Kearsley, G. (2005), Distance Education: A Systems View, Wadsworth Publishing Company.

Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. M. Newmann (Ed.), Student engagement and achievement in American secondary schools (pp. 11 – 39). New York: Teachers College Press.

Psaromiligkos, Y., Retalis, S., (2004). An approach for managing the evolution of web-based educational applications. International Journal of Web Engineering and Technology, Vol. 1. No. 2, pp.148–167.

Raptis A., (2006). Learning and Teaching in the Information Age. (Volume A), Athens

Seidman A., (2005). College student retention: Formula for student success. Greenwood Publishing Group, 2005.

Smith, P. L., & Ragan, T. J. (2005).Instructional design (3rd ed.). Hoboken, NJ: Wiley.

Standards Association of Australia. (1999).Risk analysis. Strathfield: Standards Association of Australia.

Timmermans A. C., de Wolf I. F., Bosker R. J., Doolaard S., (2015). Risk-based educational accountability in Dutch primary education.

Tinto V., (2012). Leaving college: Rethinking the causes and cures of student attrition. University of hicago Press, 2012.

Willms, J. D. (2003). Student engagement at school: A sense of belonging and participation. Paris: Organisation for Economic Co-Operation and Development.