For cited: *Kaban, A. & Bilen, Ö. (2019). A Review on Learning Analytics Researches. Journal of Current Researches on Social Sciences, 9 (1), 177-194.*

Journal of Current Researches on Social Sciences (JoCReSS)
www.stracademy.org/jocress
ISSN: 2547-9644

Year:2019
Volume: 9
Issue: 1

Review/Derleme

Crossref doi: 10.26579/jocress-9.1.11

A Review on Learning Analytics Researches

Abdullatif KABAN[1] & Ömer BİLEN[2]

Keywords
Learning analytics, ERIC, content analysis.

Article History
Received
3 Mar, 2019
Accepted
18 Mar, 2019

Abstract
In this study, which examines the research of learning analytics, it is aimed to examine the researches in ERIC database in terms of research methods and to present general trends. It was determined that 97 out of 286 articles obtained as a result of searching and filtering for this purpose were related to the subject of learning analytics and general trends were revealed through these articles. The articles have been analyzed by the content analysis method and the results show that the articles were published mostly in JLA, the articles were produced in the USA and grounded theory from qualitative methods were mostly used in articles. In the studies, LMS log data was used as a data collection tool. In the selection of the sample, the whole universe was preferred and data were collected from the students at the higher education level. These results are believed to be beneficial to the researchers who will work on learning analytics.

1. Introduction

The effort to improve the quality of education and the preliminary identification of potential problems in education led to the showing up concept of learning analysis in education. While the use of learning analytics in education is relatively new, the accurate definition of this term is still being improved (Fynn, 2016). The concept of learning analytics in the literature is often used with the concepts of academic analytics, predictive analytics, social learning analytics and educational data mining (Jayaprakash, Moody, Lauría, Regan, & Baron, 2014). Key words highlighted in the definitions of the learning analytics are can be line up as 'data collection from learning environments', 'analysis by various methods', 'use of the results for classification and estimation purposes' (Shum & Crick, 2016). Learning analytics is defined as the process of measuring, collecting, analyzing and reporting the data of learners and their contexts in order to understand and optimize the environments they are learning and experiencing (Long & Siemens, 2011). In another definition, it is mentioned that the learning analytics implements knowledge management, sociology, psychology, statistics, machine learning and data mining techniques to analyze data collected during teaching and learning, education management and services (Bienkowski, Feng, & Means, 2012).

[1] Corresponding Author. ORCID: 0000-0003-4465-3145. Asst. Prof. Dr., Bayburt University, akaban@bayburt.edu.tr
[2] ORCID: 0000-0002-6964-0781. Asst. Prof. Dr., Bayburt University, omerbilen@bayburt.edu.tr

The emergence of online learning and the accumulation of information stored in systems over time has led to the showing up concept of big data. The basis of the big data concept is the searching from the large data, generating cross-references between them and accumulation of a large number of data (boyd & Crawford, 2012). Using a variety of sources, Kitchin (2013) sorts the features of the big data concept in detail. But; it is possible to summarize the characteristics of the big data concerned with the development of the learning analytics as large volume, high speed, relationship and flexibility (Kitchin, 2013). The basic principle of learning analytics is to draw conclusions from large data sets to make a decision.

The goals of learning analytics can be often divided into four areas of results, including prediction, personalization, intervention, and information visualization (Fynn, 2016). Particularly, coming across often in the results of the studies (Colthorpe, Zimbardi, Ainscough, & Anderson, 2015; Grant, 2012; Jayaprakash, Moody, Lauría, Regan, & Baron, 2014; Pardos, Baker, San Pedro, Gowda, & Gowda, 2014; Schneider & Pea, 2015) in which the academic performance of the student can be predicted by using the existing data. The studies show that LMS log data such as sign-in frequency, site interaction, tempo, homework grades predict success (Smith, Lange, & Huston, 2012). In addition, the amount of data used in the studies has been expanded with obtain data from MOOCs and face to face environments (Merceron, Blikstein, & Siemens, 2015). Besides, it is also possible to find out results in studies that the use of learning analytics models can correlate between existing data and student achievement (Lowes, Lin, & Kinghorn, 2015; Smith, Lange, & Huston, 2012). Since it is determined that there is a positive relation between some data about learning and the final exam scores of the course (Andergassen, Mödritscher, & Neumann, 2014), the performances of the students can be predicted beforehand. Higher education institutions know that the learning analytics will make it easier to understand the participation and performance of the learners, therefore it will increase the student attendance to the school (Slade & Galpin, 2012). Learning analytics play a major role in understanding different variables that affect learning, student achievement and permanence (Prinsloo & Slade, 2013). It is also known that learning analytics are beneficial in terms of curriculum monitoring and development (Olmos & Corrin, 2012). Large data sets that accumulate in higher education institutions are analyzed to provide better information for strategic planning, decision making and pedagogical research (Ice, et al., 2012).

Long and Siemens (2011) covered the scope of the analysis at five levels: course, program, institution, region and national / international level. The first two levels are assessed as learning analytics and the other three levels are evaluated as academic analytic. Present learning analytics focuses primarily on the improvement of knowledge and skills as a result of completion of predefined tasks and activities (Chen & Zhang, 2016). Academic analytics is based on the extraction of data from one or more systems, such as the Content Management System or the Student Information System (Campbell, DeBlois, & Oblinger, 2007). But the focus of academic analytics is to improve organizational level decision-making. The use of learning analytics term is mostly used interchangeably with academic analytics, but is more specifically used to describe the use of data and models to predict

students' learning progress and performance (Leece, 2013). Learning analytics differs from academic analytics in that it analyzes the data about students to understand and optimize the learning and learning environment (Long & Siemens, 2011).

Learning analytics is a new discipline was emerged and therefore benefits from new tools and methodological approaches (Aghababyan, Martin, Janisiewicz, & Close, 2016). The development of the learning analysis in relation to related research fields such as machine learning, educational data mining, learning sciences, learning psychology and statistics has a vital value (Siemens, 2012). Learning analytics and educational data mining have many similar aims, interests and characteristics, but also differ in their technological, ideological and methodological orientations (Siemens & Baker, 2012). It is possible to find different types of analysis in the learning analytics researches such as blog analysis, sensitivity analysis, social network analysis and automatic discourse analysis. In the study of blog analysis, it was seen that the learning analysis is used for classification purposes and the authors are divided into two groups as 'entering' and 'commentator' (Pursel & Xie, 2014). Sentiment analysis, also known as opinion mining, is the use of text analysis and natural language processing to determine the attitude of a person on a subject or to reveal the general contextual intensity of a document (Kagklis, Karatrantou, Tantoula, Panagiotakopoulos, & Verykios, 2015). Social network analysis can be used to explore network learning processes by analyzing the properties of connections, what roles people take in learning relationships, and the importance of specific network occurrences (Haythornthwaite & de Laat, 2010). Social network analysis can help people understand how they develop and maintain relationships to support learning (De Laat & Prinsen, 2014).

The fact that a large amount of data is required for the implementation of learning analytics applications, and the increasing number of data, have brought about some problems. There are some ethical issues are taking the lead such as secrecy, surveillance, the nature of educational evidence (Khalil & Ebner, 2016; Prinsloo, Archer, Barnes, Chetty, & van Zyl, 2015). It is a source of concern for instructors or students that it is required techniques and methods from different disciplines such as software engineering, human-computer interaction, computer graphics, educational design and psychology for designing, validating and implementing learning analytics tools (Martinez-Maldonado, et al., 2015). Other concerns raised by researchers and practitioners in the field of learning analytics are transparency, openness, respect, user control, approval, access and accountability (Sclater, 2016). Confidentiality and ethics are important not only for learning analytics but also for analytics and big data issues in general (Steiner, Kickmeier-Rust, & Albert, 2016). The ethical issues related to learning analytics are transparency in data use, anonymization of people, ownership of data, accessibility and accuracy of results, and protection of data sets (Khalil & Ebner, 2015). Kay et al. (2012) listed the four basic principles of good analytical practice in terms of ethics:

- Clarity: Definition of purpose, scope and limits.
- Comfort and care: Consider the interests and feelings of data

- Choice and consent: Information and opportunity to opt out or opt-in.
- Conclusion and complaint: To accept the possibility of unforeseen consequences and mechanisms for complaint.

Research studies related to learning analytics are mostly possible to find them in the literature. Avella, Kebritchi, Nunn, and Kanai (2016), investigated the research that examined the methods, benefits, and difficulties of learning analytics and found 10 methods, 16 benefits, and 18 difficulty articles related to the topic from 112 articles. The methods of learning analytics are classified as visual data analysis techniques, social network analysis, semantic and prediction, clustering, relationship mining. Among the advantages of learning analytics are targeted lesson proposals, curriculum development, student learning outcomes, behavior and process, personalized learning, improved trainer performance, post-training employment opportunities, and research in the field of education. The challenges included data collection, evaluation, analysis problems, lack of connection with learning sciences, optimization of learning environments, ethics and confidentiality.

Sin and Muthu (2015), have specified 90 articles as a sample from the studies between 2011 and 2013 on learning analytics and educational data mining. They talk about three trends, introducing the concept and development of the area of learning analytics to higher education and e-learning, technical development of learning analytics framework and tools and use of learning analysis in social learning in 45 studies related to learning analytics.

Ihantola et al. (2015), examined the research conducted between 2005 and 2015 and reached 76 articles related to learning analytics and educational data mining. The articles were mostly concentrated on simple metric analysis made in a single institution and in a single course. It has been concluded that the articles were mostly focused on simple metric analysis made in a single institution and in a single course. Experiments on improving the existing studies are reported in the research.

Ochoa, Suthers, Verbert, and Duval (2014), analyzed the publications at the Third Learning Analytics and Knowledge Conference (LAK 2013) and identified five main topics: visualization, behavioral analysis, social learning analysis, learning analytics for MOOCs, and learning analytics topics.

Papamitsiou and Economides (2014), examined 40 case studies articles done between 2008 and 2013, related to learning analytics and educational data mining. It was seen that practices have often conducted within a virtual learning environment or learning management systems. Other popular topics for study include cognitive tutorials, computer-based and web-based environments, mobile settings and MOOCs, and social learning platforms.

When the researches are examined, it has been seen that there are very few articles that examine studies on learning analytics. There is a need to extract a general framework by examining the articles related to learning analytics with many aspects. This study aims to reveal general trends by examining the studies of learning analytics which were indexed by the ERIC database and published until

2017 for resolve the need. For this purpose, the following research questions have been sought:

- In which journals were learning analytics studies mainly published?
- In which countries were learning analytics studies mainly produced?
- Which research methodologies were commonly used in learning analytics studies?
- Which research designs were commonly used in learning analytics studies?
- How have sample properties changed in learning analytics studies?
- Which data collection tools were commonly used in learning analytics studies?
- What topics were commonly researched in learning analytics studies?

2. Methodology

Content analysis methodology was used to analyze each article in this study, in which studies of learning analytics indexed by the ERIC database are examined. Cohen, Manion, and Morrison (2007) emphasized that the content analysis method is a research technique that consists of organizing, classifying, comparing and extracting theoretical results from texts. In this study, the content analysis method is preferred because of the way that the readers understand the format by combining them together with the similar data within the framework of the specific concepts and themes (Büyüköztürk, Kılıç Çakmak, Akgün, Karadeniz, & Demirel, 2009).

2.1. The Study Dataset

The study dataset consisted of 286 articles published until 2017 and indexed in the ERIC database. In the study, the ERIC database was searched using the keywords "learning analytics", all the articles listed were collected by filtering back all years starting from 2016. Filtering back to old years continued until it was confident that there was no relevant article. 189 of the resulting articles were excluded from the assessment due to lack of relation of learning analytics studies. The review was conducted on 97 articles related to the subject. The numbers of the articles examined in the study according to years are shown in Table 1.

Table 1: Distribution of article sizes by years

Year	In Scope	Out of Scope	Total
2016	38	30	68
2015	23	34	57
2014	27	30	57
2013	2	8	10
2012	7	11	18
2011-1978	-	76	76
Total	97	189	286

2.2. Data Analysis Methods

The collected articles were analyzed by content analysis and the data obtained were analyzed using descriptive statistical methods. The collected data are recorded regularly in a Microsoft Excel file and the frequencies corresponding to the response of each research question are found and the percentages are calculated based on these frequencies. The results are presented as charts and graphs.

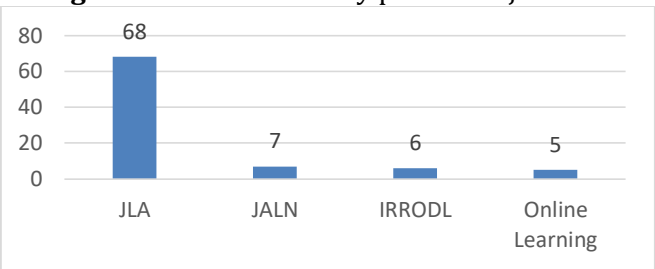
3. Findings

The data obtained from studies on ERIC database and related to learning analytics have been analyzed within the framework of research questions. The findings are presented below according to the research questions.

3.1. Most Commonly Published Journals of Learning Analytics Studies

97 articles related to the subject from the 286 articles obtained in the scope of the study were examined and it was determined that these studies appeared after 2012. Figure 1 shows the journals which number of articles is 5 and above.

Figure 1: Most commonly published journals



When looking at Figure 1, it appears that the journal with the most commonly article published in is Journal of Learning Analytics (JLA). This journal is followed by Journal of Asynchronous Learning Networks (JALN), The International Review of Research in Open and Distributed Learning and Online Learning (IRRODL) respectively. The distribution of the articles in the examined journals according to years is shown in Table 2.

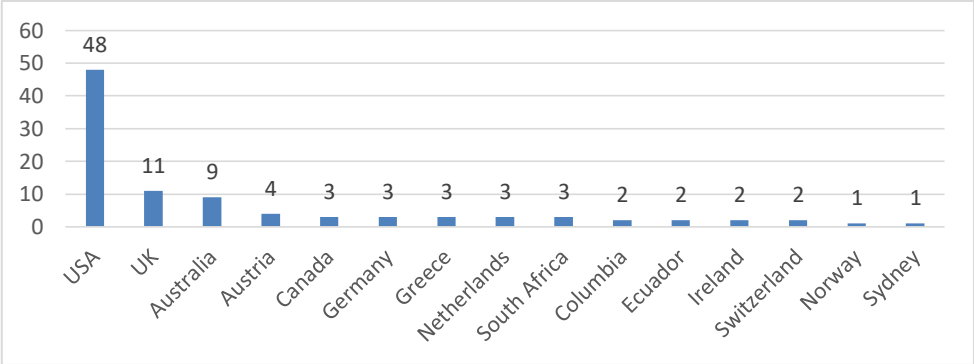
Table 2: Distribution of the articles in the journals by years

Journals	2012	2013	2014	2015	2016	Total
Journal of Learning Analytics			21	19	28	68
Journal of Asynchronous Learning Networks (JALN)	6	1				7
The International Review of Research in Open and Distributed Learning			2	2	2	6
Online Learning					5	5
Research & Practice in Assessment			2			2
Journal of Interactive Media in Education			1		1	2
Contemporary Educational Technology			1			1
Continuing Higher Education Review	1					1
Canadian Journal of Learning and Technology		1				1
European Journal of Open, Distance and e-Learning				1		1
International Journal for the Scholarship of Teaching and Learning					1	1
Journal of University Teaching & Learning Practice					1	1
Knowledge Quest				1		1
Total	7	2	27	23	38	97

3.2. Most Commonly Produced Countries of Learning Analytics Studies

The countries studied by the first authors were taken into account in order to determine the countries where learning analytics articles were produced. The numbers thus obtained are given in Figure 2. According to the graph, it has been seen that most of the articles are produced in USA. The UK, Australia and Austria are the other countries where the most of the articles on learning analytics are produced respectively.

Figure 2: Countries where learning analytics studies are produced



The distributions of the numbers and years according to the countries where their studies of learning analytics are produced are given in Table 3. When we look at Table 3, it can be said that there is generally an increase in the annual rate. However, in 2012 there is no publication except for the USA and Australia and in 2013 there is no publication except for the USA and Germany.

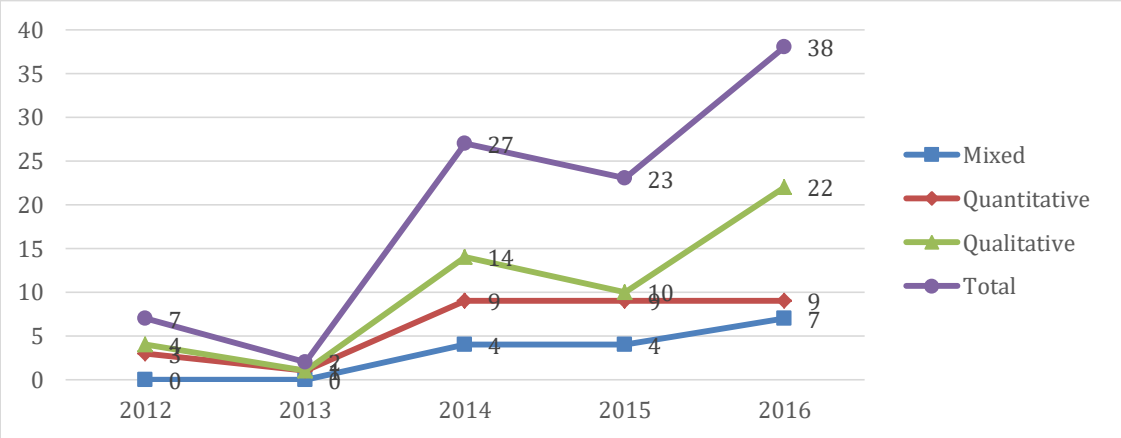
Table 3: Distribution of country numbers of learning analytics studies by years

Countries	2012	2013	2014	2015	2016	Total
USA	6	1	11	9	21	48
UK			4	3	4	11
Australia	1		1	4	3	9
Austria			1	1	2	4
Canada			1	2		3
Germany			2	1		3
Greece		1	1	1		3
Netherlands			1	1	1	3
South Africa				1	2	3
Columbia			2			2
Ecuador			2			2
Ireland			1		1	2
Switzerland					2	2
Norway					1	1
Sydney					1	1

3.3. Most Commonly Used Research Methodologies of Learning Analytics Studies

The research methodologies that are most commonly preferred in the learning analytics studies and the numbers according to years are given in Figure 3. When Figure 3 is examined, it has been seen that publishing started after 2012, relatively decreased in 2013 and 2015, and increased in 2014 and 2016 compared to the previous year. Mixed methods have begun to appear after 2014. It has been seen that qualitative methods are commonly preferred in the researches and quantitative and mixed methods follow it respectively

Figure 3: Distribution of research methods preferred in learning analytics studies by years



3.4. Most Commonly Used Research Designs of Learning Analytics Studies

The research designs that are widely preferred in the articles of learning analytics are given in Table 4. When Table 4 is examined, it has been seen that the most preferred research design is grounded theory (45,36%) within qualitative studies. Correlation studies (26,80%) from quantitative research methods and variance studies (12,37%) from mixed methods follow the sequence as the most preferred designs.

Table 4: Research Patterns Commonly Preferred in Learning Analytic Studies

Research Methodologies	Research Designs	N	%
Qualitative		51	52,58%
	Grounded theory	44	45,36%
	Case study	4	4,12%
	Literature review	2	2,06%
	Phenomenological	1	1,03%
Quantitative		31	31,96%
	Correlational	26	26,80%
	Experimental	2	2,06%
	Survey research	2	2,06%
	Quasi-experimental	1	1,03%
Mixed		15	15,46%
	Triangulation	12	12,37%
	Exploratory	2	2,06%
	Design based	1	1,03%
Total		97	100,00%

3.5. Most Commonly Preferred Sample Properties of Learning Analytics Studies

The sample properties of learning analytics articles were examined in three dimensions as level, sample selection and size. When the frequency of the sample is given in Table 5 is examined, it has been seen that the most preferred sample level higher education (52%). This is followed by the level of secondary education (26%).

Table 5: Sample level frequency of learning analytics studies

Sample Level	N	%
Higher education	26	52%
Secondary education	13	26%
Graduate	5	10%
Online course	4	8%
Mixed stakeholders	1	2%
Teacher	1	2%
Total	50	100%

When Table 6, which provides common used methods of sample selection in the studies of learning analytics, is examined it has been seen that most of the studies used the whole universe (50,98%). This is followed by convenience (21.57%) and purposive (19.61%) sample selection methods.

Table 6: Sample selection methods of learning analytics studies

Sample Selection Methods	**N**	**%**
Whole universe	26	50,98%
Convenience	11	21,57%
Purposive	10	19,61%
Voluntary	2	3,92%
Random	2	3,92%
Total	51	100,00%

The distribution of sample sizes by research methodologies of learning analytical studies are given in Table 7. When the Table 7 is examined, it has been seen that the group of More than 1001 persons (23,5%) is the most preferred in the quantitative methodology studies. It is followed by the group of 101-1001 persons (21,6%) is the most preferred in the quantitative methodology studies. Also seen that the group of 11-100 persons is the most preferred in the mixed (15,7%) and qualitative (7,8%) methodology studies.

Table 7: Distribution of sample sizes by methodologies of learning analytics studies

Sample Size	Quantitative		Mixed		Qualitative		Total	
	n	%	n	%	n	%	n	%
1-10	-	-	1	2,0	1	2,0	2	3,9
11-100	6	11,8	8	15,7	4	7,8	18	35,1
101-1000	11	21,6	4	7,8	1	2,0	16	31,4
More than 1001	12	23,5	2	3,9	1	2,0	15	29,4

3.6. Most Commonly Preferred Data Collection Tools of Learning Analytics Studies

Data collection tools commonly preferred in learning analytics studies are given in Table 8. When the findings are examined, it has been seen that system log data (53,2%) is commonly preferred as a data collection tool in quantitative studies. Surveys (12,8%) in mixed studies and system log data (4,3%) and interview (4,3%) in qualitative studies are commonly preferred. When looking at the total numbers, it has been seen that the data collection tool which is commonly preferred is the system log data (63,8%), followed by the survey (14,9) and interview (8,5%).

Table 8: Distribution of data collection tools by methodologies of learning analytics studies

Data Collection Tool	Quantitative		Mixed		Qualitative		Total	
	n	%	n	%	n	%	n	%
System log data	25	53,2	3	6,4	2	4,3	30	63,8
Survey	1	2,1	6	12,8	-	-	7	14,9
Interview	-	-	2	4,3	2	4,3	4	8,5
Observation	-	-	2	4,3	1	2,1	3	6,4
Text scanning	-	-	-	-	1	2,1	1	2,1
Literature search	-	-	-	-	1	2,1	1	2,1
Exam score	1	2,1	-	-	-	-	1	2,1

3.7. Most Commonly Preferred Subjects of Learning Analytics Studies

When the articles are examined in terms of subjects related to learning analytics, it has been seen that the number of quantitative studies is higher. Predictive models and social learning analytics are emphasized, although many different topics are mentioned. Learning analytics subjects commonly preferred in learning analytics studies are given in Table 9.

Table 9: Distribution of learning analytics subjects by methodologies of learning analytics studies

Learning Analytics Subjects	**Mixed**	**Quantitative**	**Qualitative**	**Total**
Predictive models	2	2	1	5
Social learning analytics		2	1	3
Logistic regression model		2		2
Multimodal learning analytics	1		1	2
Natural language processing algorithms	1		1	2
Others	4	15	12	31
Total	8	21	16	45

The topics not listed in Table 9 and only mentioned once are listed below.

- De-identification technics: anonymization, masking, blurring,
- Action analytics, big data, top-down/bottom-up
- Agency-driven, choice-based analytics
- Bottom up
- Business intelligence techniques
- Cluster analysis
- Coherence analysis
- Course signals (cs) project
- Data protection framework
- Discourse-centric learning analytics

- Dispositional learning analytic
- Educational content analytics
- Educational data science
- Evidence-based learning analytics
- Experimentation strategies in design model
- K-means clustering algorithm
- Learning analytics design space model
- Learning analytics framework
- Learning factor models
- Measurement model
- Model tracing approach, novice-expert overlay models, sub group discovery method
- Productive persistence
- Reflective writing analytics
- Scientometric method of main path analysis
- Structural equation model, learning analytics for performance and action
- Structural topic model
- Student tuning model
- Text mining, sentiment analysis
- Video analytics
- Visualization
- Web usage mining

4. Discussion and Conclusion

In this study, the studies indexed by the ERIC database and related to learning analytics were examined in terms of the published journals, produced countries, used research methods and designs, preferred sample properties, data collection tools and learning analytics subjects.

In this study, starting from 2016, it was backward scanned and the articles reached were examined. In the ERIC database, it was seen that 97 out of 286 articles resulting from the scan with the keywords "learning analytics" were related to learning analytics and the studies started to show up after 2012. The pre-existing learning analytics studies have gained momentum thanks to the been mentioned of this subject in the 2012 NIMC Horizon Report (Johnson, Adams, & Cummins, 2012) and this have provided researchers tend to this subject.

When the published journals are examined, it has been seen that JLA has a big share. The journal, which started to be published since 2014, has increased its preference rate due to its dedicated service for learning analytics.

The countries in which the first authors work has been taken into account in order to understand which countries the articles were published from. As a result of the review, it was determined that the most articles was published from the USA. The reason for this may be that the first learning analytics researches has come from the USA. This is followed by UK and Australia. Because data processing and analysis are the basis of learning analytics researches, it is natural that the countries with distance learning universities are closely interested in this issue.

It was concluded that qualitative methods were widely preferred in Learning analytics researches. As learning analytics is a field of study that can be considered as new, the attempt to construct theoretical foundations have led to the increase of such studies. When the used research designs are examined, it has been concluded that grounded theory was commonly preferred. This result supports the reason why qualitative studies are preferred. The reason for preferring correlational studies within quantitative methods can be explained by the fact that learning analytics studies have tried to put forward the predictability level of the relation between some data of the student and academic success.

When the articles were examined in terms of sample characteristics, it was concluded that the higher education level was mostly preferred. Distance education institutions offer learning environments through digital systems to students and all kinds of movements of students are recorded. This makes it easier to carry out learning analytics studies with this data which is already readily available in distance education institutions. When we look at sample selection methods, mostly all of the universe was used, followed by convenience and purposive sampling methods. Because it is possible to reach the full range of students enrolled in distance education systems, learning analytics have the opportunity to use the entire universe in their work. Because it is possible to reach the full data of students enrolled in distance education systems, there is an opportunity to use the entire universe in learning analytics studies. When the balanced distribution of sample numbers is taken into consideration, it is considered that the number of students enrolled in the system affects the sample numbers.

When the data collection tools commonly used in learning analytics studies are examined, it is seen that system log data are used, especially in quantitative studies. System log data consist of students' behavior on the learning management systems. They are tried to have an idea about the students' academic achievements of the students by considering the time and duration of the students' participation in the course, their interactions with the content, etc. Smith et al. (2012) also came to the conclusion that academic success can be predict by LMS log data such as sign-on frequency, site interaction, tempo, homework grades. The healthiest data to be used in learning analytics studies is obtained from these records. The healthiest data to be used in 'learning analytics' studies can be obtained from these log data.

When the studies are examined, it is seen that the most preferred learning analytics subject is "predictive models". Learning analytic studies are mainly aimed at predicting the success of the student and then make improvements in teaching services for the learners by taking the necessary precautions. Therefore, it is considered that the studies of learning analytics based on prediction will increase gradually. The reason why the study subjects are so diverse and it is not yet fully focused on a subject is that the field of learning analytics is both young and has a wide range of work areas. It is anticipated that the learning analytics subjects that will be preferred will be concentrated at certain points in the future.

The following suggestions are made in light of the results;

- Qualitative methods are commonly used in learning analytic studies. The number of quantitative researches should be increased by considering these studies made for the establishment of theoretical foundations and the validity of the theories presented should be tested.
- Distance education is now widely used in many countries of the world. Distance education institutions can increase 'learning analytics' studies using stored data in their systems so learning analytics studies can be spread all over the world. Thus, the regional differences of the theoretical foundations founded by the studies of learning analytics can also be revealed.
- Massive open online courses that continue to spread with distance education have made it possible to reach participants at different levels. Selection of the sample to be used in the 'learning analytics' studies from these courses will provide data diversity and different outcomes emerged will be used to improve the quality of education.
- For most of the studies, the sample was selected from students of a single institution or a single course. Data diversity in the studies can be achieved by selecting sample from different universities or courses
- Only ERIC is used in this study. Similar studies can be done using other databases relevant to education. The results obtained will be helpful to other work.

References

Aghababyan, A., Martin, T., Janisiewicz, P., & Close, K. (2016). Microgenetic learning analytics methods: Workshop report. *Journal of Learning Analytics, 3*(3), 96-114. doi:http://dx.doi.org/10.18608/jla.2016.33.6

Andergassen, M., Mödritscher, F., & Neumann, G. (2014). Practice and Repetition during Exam Preparation in Blended Learning Courses: Correlations with Learning Results. *Journal of Learning Analytics, 1*(1), 48-74. doi:http://dx.doi.org/10.18608/jla.2014.11.4

Avella, J. T., Kebritchi, M., Nunn, S., & Kanai, T. (2016). Learning Analytics Methods, Benefits, and Challenges in Higher Education: A Systematic Literature Review. *Online Learning, 20*(2), 13-29. doi:http://dx.doi.org/10.24059/olj.v20i2.790

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief.* U.S. Department of Education, Office of Educational Technology. Retrieved July 12, 2017, from https://tech.ed.gov/wp-content/uploads/2014/03/edm-la-brief.pdf

boyd, d., & Crawford, K. (2012). Critical Questions for Big Data. *Information, Communication & Society, 15*(5), 662-679. doi:http://dx.doi.org/10.1080/1369118X.2012.678878

Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2009). *Bilimsel Araştırma Yöntemleri.* Ankara: Pegem Akademi.

Campbell, J. P., DeBlois, P. B., & Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review, 42*(4), 40-57. Retrieved from http://www.educause.edu/ir/library/pdf/erm0742.pdf

Chen, B., & Zhang, J. (2016). Analytics for Knowledge Creation: Towards Epistemic Agency and Design-Mode Thinking. *Journal of Learning Analytics, 3*(2), 139-163. doi:http://dx.doi.org/10.18608/jla.2016.32.7

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education.* Taylor & Francis e-Library.

Colthorpe, K., Zimbardi, K., Ainscough, L., & Anderson, S. (2015). Know Thy Student! Combining Learning Analytics and Critical Reflections to Increase Understanding of Students' Self-Regulated Learning in an Authentic Setting. *Journal of Learning Analytics, 2*(1), 134-155. doi:http://dx.doi.org/10.18608/jla.2015.21.7

De Laat, M., & Prinsen, F. R. (2014). Social Learning Analytics: Navigating the Changing Settings of Higher Education. *Research & Practice in Assessment, 9*(Winter), 51-60. Retrieved from http://www.rpajournal.com/dev/wp-content/uploads/2014/10/A5.pdf

Fynn, A. (2016). Ethical Considerations in the Practical Application of the Unisa Socio-Critical Model of Student Success. *International Review of Research in Open and Distributed Learning, 17*(6), 206-220. doi:http://dx.doi.org/10.19173/irrodl.v17i6.2812

Grant, M. R. (2012). Data-Driven Online Course Design and Effective Practices. *Continuing Higher Education Review, 76*, 183-192.

Haythornthwaite, C., & de Laat, M. (2010). Social networks and learning networks: Using social network perspectives to understand social learning. *Proceedings of the 7th International Conference on Networked Learning*, (pp. 183-190).

Ice, P., Díaz, S., Swan, K., Burgess, M., Sharkey, M., Sherrill, J., . . . Okimoto, H. (2012). The PAR Framework Proof of Concept: Initial Findings from a Multi-Institutional Analysis of Federated Postsecondary Data. *Journal of Asynchronous Learning Networks, 16*(3), 63-86. doi:http://dx.doi.org/10.24059/olj.v16i3.277

Ihantola, P., Vihavainen, A., Ahadi, A., Butler, M., Butler, M., Börstler, J., . . . Toll, D. (2015). Educational Data Mining and Learning Analytics in Programming: Literature Review and Case Studies. *Proceedings of the 2015 ITiCSE on Working Group Reports*, (pp. 41-63). New York, NY, USA.

Jayaprakash, S. M., Moody, E. W., Lauría, E. J., Regan, J. R., & Baron, J. D. (2014). Early Alert of Academically At-Risk Students: An Open Source Analytics Initiative. *Journal of Learning Analytics, 1*(1), 6-47. doi:http://dx.doi.org/10.18608/jla.2014.11.3

Johnson, L., Adams, S., & Cummins, M. (2012). *The NMC Horizon Report: 2012 Higher Education Edition.* Austin, Texas: The New Media Consortium.

Kagklis, V., Karatrantou, A., Tantoula, M., Panagiotakopoulos, C. T., & Verykios, V. S. (2015). A Learning Analytics Methodology For Detecting Sentiment In Student Fora: A Case Study In Distance Education. *European Journal of Open, Distance and e-Learning, 18*(2), 74-94. doi:https://doi.org/10.1515/eurodl-2015-0014

Kay, D., Korn, N., & Oppenheim, C. (2012). Legal, risk and ethical aspects of analytics in higher education. *JISC CETIS Analytics Series, 1*(6).

Khalil, M., & Ebner, M. (2015). Learning Analytics: Principles and Constraints. *In Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications*, (pp. 1326-1336).

Khalil, M., & Ebner, M. (2016). De-Identification in Learning Analytics. *Journal of Learning Analytics, 3*(1), 129–138. doi:http://dx.doi.org/10.18608/jla.2016.31.8

Kitchin, R. (2013). Big Data and Human Geography: Opportunities, Challenges and Risks. *Dialogues in Human Geography, 3*(3), 262-267. doi:https://doi.org/10.1177/2043820613513388

Leece, R. (2013). Analytics: An exploration of the nomenclature in the student experience. *Proceedings of the 16th International First Year in Higher Education Conference (FYHE13).* Te Papa Tongarewa, Wellington, New Zealand.

Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *Educause Review, 46*(5), 31-40.

Lowes, S., Lin, P., & Kinghorn, B. (2015). Exploring the link between online behaviours and course performance in asynchronous online high school courses. *Journal of Learning Analytics, 2*(2), 169-194. doi:http://dx.doi.org/10.18608/jla.2015.22.13

Martinez-Maldonado, R., Pardo, A., Mirriahi, N., Yacef, K., Kay, J., & Clayphan, A. (2015). LATUX: An iterative workflow for designing, validating, and deploying learning analytics visualizations. *Journal of Learning Analytics, 2*(3), 9-39. doi:http://dx.doi.org/10.18608/jla.2015.23.3

Merceron, A., Blikstein, P., & Siemens, G. (2015). Learning Analytics: From Big Data to Meaningful Data. *Journal of Learning Analytics, 2*(3), 4-8. doi:http://dx.doi.org/10.18608/jla.2015.23.2

Ochoa, X., Suthers, D., Verbert, K., & Duval, E. (2014). Analysis and Reflections on the Third Learning Analytics and Knowledge Conference (LAK 2013). *Journal of Learning Analytics, 1*(2), 5-22. doi:http://dx.doi.org/10.18608/jla.2014.12.2

Olmos, M., & Corrin, L. (2012). Learning Analytics: A Case Study of the Process of Design of Visualizations. *Journal of Asynchronous Learning Networks, 16*(3), 39-49. doi:http://dx.doi.org/10.24059/olj.v16i3.273

Papamitsiou, Z., & Economides, A. (2014). Learning Analytics and Educational Data Mining in Practice: A Systematic Literature Review of Empirical Evidence. *Educational Technology & Society, 17*(4), 49-64.

Pardos, Z. A., Baker, R. S., San Pedro, M. O., Gowda, S. M., & Gowda, S. M. (2014). Affective States and State Tests: Investigating How Affect and Engagement during the School Year Predict End-of-Year Learning Outcomes. *Journal of Learning Analytics, 1*(1), 107-128. doi:http://dx.doi.org/10.18608/jla.2014.11.6

Prinsloo, P., & Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning analytics. *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, (pp. 240-244). doi:http://dx.doi.org/10.1145/2460296.2460344

Prinsloo, P., Archer, E., Barnes, G., Chetty, Y., & van Zyl, D. (2015). Big(ger) Data as Better Data in Open Distance Learning. *International Review of Research in Open and Distributed Learning, 16*(1), 284-306. doi:http://dx.doi.org/10.19173/irrodl.v16i1.1948

Pursel, B. K., & Xie, H. (2014). Patterns and Pedagogy: Exploring Student Blog Use in Higher Education. *Contemporary Educational Technology, 5*(2), 96-109.

Schneider, B., & Pea, R. (2015). Does seeing one another's gaze affect group dialogue A computational approach. *Journal of Learning Analytics, 2*(2), 107-133. doi:http://dx.doi.org/10.18608/jla.2015.22.9

Sclater, N. (2016). Developing a Code of Practice for Learning Analytics. *Journal of Learning Analytics, 3*(1), 16-42. doi:http://dx.doi.org/10.18608/jla.2016.31.3

Shum, S. B., & Crick, R. D. (2016). Learning Analytics for 21st Century Competencies. *Journal of Learning Analytics, 3*(2), 6-21. doi:http://dx.doi.org/10.18608/jla.2016.32.2

Siemens, G. (2012). Learning Analytics: Envisioning a Research Discipline and a Domain of Practice. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*, (pp. 4-8). doi:http://dx.doi.org/10.1145/2330601.2330605

Siemens, G., & Baker, R. S. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, (pp. 252-254). New York, NY, USA.

Sin, K., & Muthu, L. (2015). Application of Big Data in Education Data Mining and Learning Analytics – A Literature Review. *ICTACT Journal on Soft Computing, 5*(4), 1035-1049.

Slade, S., & Galpin, F. (2012). Learning Analytics and Higher Education: Ethical Perspectives. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK12)*, (pp. 16-17). Vancouver, BC, Canada.

Smith, V. C., Lange, A., & Huston, D. R. (2012). Predictive Modeling to Forecast Student Outcomes and Drive Effective Interventions in Online Community College Courses. *Journal of Asynchronous Learning Networks, 16*(3), 51-61.

Steiner, C. M., Kickmeier-Rust, M. D., & Albert, D. (2016). LEA in private: A privacy and data protection framework for a learning analytics toolbox. *Journal of Learning Analytics, 3*(1), 66-90. doi:http://dx.doi.org/10.18608/jla.2016.31.5