



Faculty of Cognitive Sciences and Human Development

A Bibliometric Study on Learning Analytics

Nurul Athira Nahar Binti Ahmad Nasir

**Cognitive Science
2022**

KMK3104 FINAL YEAR PROJECT 2

I declare this Project/Thesis is classified as (Please tick (√)):

- CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*
- RESTRICTED** (Contains restricted information as specified by the organisation where research was done)*
- OPEN ACCESS**

I declare this Project/Thesis is to be submitted to the Centre for Academic Information Services (CAIS) and uploaded into UNIMAS Institutional Repository (UNIMAS IR) (Please tick (√)):

- YES**
- NO**

Validation of Project/Thesis

I hereby duly affirmed with free consent and willingness declared that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abide interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic and research purposes only and not for other purposes.
- The Centre for Academic Information Services has the lawful right to digitize the content to be uploaded into Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis if required for use by other parties for academic purposes or by other Higher Learning Institutes.
- No dispute or any claim shall arise from the student himself / herself neither a third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student himself/herself without first obtaining approval from UNIMAS.

Student's signature: 
Date: 26 JANUARY 2022

Supervisor's signature: 
Date: 27 Jan 2022

Current Address:

Notes: * If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the date of restriction indicated, and the reasons for the confidentiality and restriction.

KMK3104 FINAL YEAR PROJECT 2

A BIBLIOMETRICS STUDY ON LEARNING ANALYTICS

NURUL ATHIRA NAHAR BINTI AHMAD NASIR

This project is submitted
in partial fulfillment of the requirements for
a Bachelor of Science with Honours
(Cognitive Science)

Faculty of Cognitive Sciences and Human Development
UNIVERSITI MALAYSIA SARAWAK
(2022)

KMK3104 FINAL YEAR PROJECT 2

The project entitled 'A Bibliometric Study on Learning Analytics' was prepared by Nurul Athira Nahar Binti Ahmad Nasir and submitted to the Faculty of Cognitive Sciences and Human Development in partial fulfillment of the requirements for a Bachelor of Science with Honours (Cognitive Science).

Received for examination by:



(PROFESSOR DR CHEN CHWEN JEN)

Date: 27 Jan 2022

Grade A

KMK3104 FINAL YEAR PROJECT 2

ACKNOWLEDGEMENTS

Thank you for my friends', families', and lecturers' support and encouragement. I would like to express my appreciation to a few individuals in particular. Especially to Professor Dr Chen Chwen Jen, my supervisor, I appreciate your willingness to assist and lead me on this path, as well as your support and trust throughout the process. Thank you, Dr Tan Kock Wah, for assisting me in planning my final year project with an outstanding supervisor. I appreciate your contribution towards this study.

Finally, for my family, our prayers have been heard. Thank you for your kind words of encouragement.

CONTENT

CHAPTER 1 INTRODUCTION

1.0 Introduction	12-14
1.1 Problem Statement	14
1.2 Research Objectives	15
1.3 Research Questions	15
1.4 Definitions and Concepts	15-16
1.5 Significance of the Study	17
1.6 Scope of the Study	17

CHAPTER 2 LITERATURE REVIEW

2.0 Introduction	18
2.1 Bibliometrics Analysis	18
2.1.1 Analysis of the Citations	18-19
2.2 Learning Analytics	19-20
2.3 Learning Analytics Theoretical Model	21
2.4 Learning Analytics for Educational Purposes	22
2.5 Previous Studies	23

CHAPTER 3 METHODOLOGY 24-26

CHAPTER 4 RESULTS

4.0 Findings from the Research and Analysis	27
4.1 Present condition of publication in learning analytics	27
4.1.1 Year of publications	28-30
4.1.2 Types of papers	30-31
4.1.3 Publication by source titles	31-34
4.1.4 Type of source	34-35
4.1.5 Publication by country	35-36
4.1.6 Publication by the institution	36-38
4.1.7 Languages of documents	38-39
4.1.8 Subject area	39-41
4.2 Citations Trends in learning analytics	41-45

CHAPTER 5 DISCUSSION & CONCLUSION

5.0 Introduction	46
5.1 Discussion	46
5.1.1 Present condition of learning analytics publications	46-47

KMK3104 FINAL YEAR PROJECT 2

5.1.2 Recent citation trends for learning analytics publications	47
5.2 Limitation of the Study	47-48
5.3 Future Research Recommendations	48
5.4 Conclusion	49
References	50-55

ABSTRACT

Each day, new learning analytics tools and approaches are represented in academic journals. This study will gather and analyse literature on the subject from 2011 to 2020 in order to better understand the essential foundation of Learning Analytics. To begin, citation and bibliometric analyses of 1871 Scopus papers revealed the top authors, affiliations, subject areas, languages, and sources of publication (journals and conferences). Second, VOSviewer software is utilised to examine the association between citations by country and learning analytics source titles. Learning Analytics have sparked the interest of the global population, according to the findings of this study. Australia, the United States, the United Kingdom, and Spain are leading contributors to information distribution in the field of learning analytics. The CEUR Workshop Proceedings, ACM International Conference Proceeding Series, and Computer Science Lecture Notes that consist Bioinformatics and Artificial Intelligence Notes. are the most widely used publications. Rebecca Ferguson; Wolfgang Greller and Hendrik Drachsler; as well as George Siemens and Ryan SJ D. Baker are the top authors that received high total citations for their publications on educational data analytic.

Keywords: Bibliometrics, Learning Analytics, Educational Data Analytics

ABSTRAK

Setiap hari, alat dan pendekatan analitik pembelajaran baharu diwakili dalam jurnal akademik. Kajian ini akan mengumpulkan dan menganalisis literatur mengenai subjek dari 2011 hingga 2020 untuk memahami dengan lebih baik asas penting Analitis Pembelajaran. Sebagai permulaan, petikan dan analisis bibliometrik kertas Scopus 1871 mendedahkan pengarang, gabungan, bidang subjek, bahasa, dan sumber penerbitan (jurnal dan persidangan) teratas. Kedua, perisian VOSviewer digunakan untuk mengkaji perkaitan antara petikan mengikut negara dan tajuk sumber analisis pembelajaran. Pembelajaran Analitis telah mencetuskan minat penduduk global, menurut penemuan kajian ini. Australia, Amerika Syarikat, United Kingdom dan Sepanyol merupakan penyumbang utama kepada pengedaran maklumat dalam bidang analisis pembelajaran. Prosiding Bengkel CEUR, Siri Prosiding Persidangan Antarabangsa ACM, dan Nota Kuliah Sains Komputer yang terdiri daripada Nota Bioinformatik dan Kecerdasan Buatan. merupakan penerbitan yang paling banyak digunakan. Rebecca Ferguson; Wolfgang Greller dan Hendrik Drachsler; serta George Siemens dan Ryan SJ D. Baker ialah pengarang yang menerima jumlah petikan tertinggi untuk penerbitan mereka mengenai analisis data pendidikan.

Kata kunci: Bibliometrik, Analitis Pembelajaran, Analitis Data Pendidikan

CHAPTER 1

INTRODUCTION

1.0 Introduction

Technology is assisting in the transformation of education, changing how knowledge is transmitted and the reach of teaching and learning. A student's online activity log is created every time they connect with their university, whether it's submitting tests online, logging into their virtual learning space, or going to the library. The method of using information to improve education is known as learning analytics. According to Siemens and Long (2011), they stated that the goal of learning analytics is to assess human behaviour in the learning process of students, then analyze and define it to explore new opportunities and offer advancements to stakeholders for enhancing learning, decision making, teaching and efficient organization.

Moreover, big data depends on the ability to store massive amounts of data for long periods and down to individual transactions (Picciano, 2012). The open-source platform, learning management systems, open social platforms, and numerous online tools can all be used to collect big data (Reyes, 2015). In addition, analytics is a scientific method that evaluates data to create findings and propose methods to make important decisions, and similar to data-driven decision making (Picciano, 2012). Learning analytics (LA), based on Brown (2012) study, is the practice of methodically gathering and evaluating big data samples from online sources with the intent of enhancing knowledge developments.

In the educational world, learning analytics is a growing field. In this case, learning analytics will be frequently utilized in online education in the next several years, according to experts in online learning in American higher education, to identify students' patterns of

KMK3104 FINAL YEAR PROJECT 2

behaviour and to increase students' learning and retention rates. The predictive models are used in learning analytics to deliver actionable data. Technological learning enhancement, educational data mining, visualization and data processing are all part of this multidisciplinary approach (Scheffel, Drachsler, Stoyanov, & Specht, 2014). The aim of educational data analytics is to personalize opportunities for education to the needs and abilities of each learner, such as assisting those at-risk pupils or giving advice and teaching material. However, educational data mining, on the other hand, aims to generate systematic and automated answers to students.

Other than that, academic analytics, educational data mining, and learning analytic are all strongly linked topics (Bienkowski, Feng, & Means, 2012; Elias, 2011). The purpose of educational data mining is to create and apply approaches for boosting data findings in academic contexts. It looks for trends in a huge dataset of student performance. The techniques could be used to gain a better knowledge of educational contexts and students. Next, data mining is defined by Hung, Hsu, and Rice (2012) as data analysis approaches that, when implemented, retrieve meaningful information through activities such as pattern discovering and predictive modelling. Educational data analytics is the practice of business intelligence principles and techniques to academics with the objectives of enhancing prediction and results (Campbell, De Blois, & Oblinger, 2007).

Other than higher education, learning analytics is not as well-known or used as it is in higher education itself (Drachsler & Greller, 2012). Ferguson (2012) established three different potential learning analytics users or interest organization which are educational institutions, teachers, and governments. These groups are defined as focus levels based on three drivers of learning analytics which are online learning political issues and big data. Each interest group

dictates how data is studied, including how researchers collect data, draw conclusions, act on these study results, and develop their ideas (Ferguson, 2012).

Bibliometrics are mathematical or statistical methods for evaluating the quantity and quality of published scientific literature, as well as for studying journal analysis, citation analysis, international, authorship, publication impact, research trends and national contributions in a field. Many scientific disciplines, including business, education and technology, have used bibliometrics. Since it "offers an objective tool to define, classify, and track published research," bibliometric analysis is beneficial (Walsh & Renaud, 2017). A bibliometric study examines the quantitative characteristics of scholarly publications using statistical tools (Andres, 2009; Waheed, Hassan, Aljohani & Wasif, 2018; Walsh & Renaud, 2017).

1.1 Problem Statement

Learning analytics and its likely to improve learning and teaching settings are generating a lot of excitement (Siemens, 2011). Since 2004, about 4,044 learning analytics documents, have been released through the Scopus database. Adeniji (2019) presents a deeper analysis of learning analytics and the study is based on the Scopus database. However, Adeniji (2019) study period starts from 2004 until 2018. Therefore, this study looked at scholarly publications based on the Scopus database to gather and evaluate conversation and discover knowledge structures in the Learning Analytics community from 2011 until 2020.

1.2 Research Objectives

General Objective

The purpose of the research is to explore and produce a comprehensive bibliometric analysis of Learning Analytics (LA) for education by using the Scopus database.

Specific Objective

The specific objectives are:

1. to identify the recent state of publications on learning analytics
2. to identify the existing citation patterns of publications on learning analytics

1.3 Research Questions

This research was conducted to answer these research questions:

Research question 1: What is the present condition of learning analytics publications?

Research question 2: What are the most recent citation trends for learning analytics publications?

1.4 Definition and Concepts

This part of the proposed study explains some fundamental definitions and ideas that are applied in research document review.

Bibliometrics

Bibliometrics is a branch of science that analyses bibliometric data statistically (Pritchard, 1969). Bibliometrics now encompasses a varied range of tools and methods for analysing and visualising data (Small, 1999).

Cite Scores

The citation metrics were produced using Scopus social network analysis, and they indicate the relative success of serial titles over time.

Citation Analysis

Citation analysis is a bibliometrics method for examining the link between cited and citing documents (Diodato, 1994).

NMC Horizon Report

The research collection is the world's oldest investigation of developing technology developments in university education. Over the course of five years, the series follows the implications of modern methods and approaches on university education around the world (Adeniji, 2019).

1.5 Significance of the study

The intellectual frameworks discovered in scholarly publications on Learning Analytics were investigated in this research project. The study looked at the research trends that have emerged since the inception of educational data analytics. It gives details on educational data analytic publications, tools, and strategies that can be used to help stakeholders improve the environments for research and education. The discoveries of this study provide vision into learning analytics scholarly research as well as proof of its utilisation by stakeholders in learning analytics discourse groups.

1.6 Scope of the study

This study focuses on exploring and producing a comprehensive bibliometric analysis of Learning Analytics (LA) for education by using the Scopus database. The data of publications from the year 2011-2020 is taken on 1st December 2021. This study will be conducted based on the Scopus database only. This research also reviewed these research questions (RQs): Research question 1: What is the present condition of learning analytics publications? Research question 2: What are the most recent citation trends for learning analytics publications?

The result of this study is looking at the publishing trends (language, document type, countries etc) and the total citations of the publications that involve in the evolution of Learning Analytics (LA).

CHAPTER 2

LITERATURE REVIEW

2.0 Introduction

The exploratory studies on learning analytics and bibliometrics for knowledge organisations is summarised in this chapter. The chapter also discusses learning analytics for educational purposes, conceptual frameworks, and previous learning analytics research.

2.1 Bibliometrics Analysis

Library and information science professionals employ bibliometric tools and techniques to investigate information flows and communication processes, as well as to better comprehend knowledge management and distribution (Adeniji, 2019). Bibliometric approaches are now utilized to evaluate research development, choose publications for academic libraries, and forecast a topic's or profession's future. This research explained the comprehensive literature features of educational progress in terms of educational data analytics patterns. The paper also includes a timeline of Learning Analytics' progress from 2011 until 2020.

2.1.1 Analysis of the Citations

The analysis of the citations is a method of counting publications, finding peers, tracking public revolution, and disseminating information in a database's core journals. The quantity of citations in various writings is used to assess a writer's, paper's, or journal's relative importance or consequences. Evidence of knowledge diffusion is an important part of topic development; this research looked at how learning analytics publications and forums are associated with academic research. Educational libraries provide entry with a vast amount of

information that may be used for study and writing. The *h*-index also known as Hirsch index is the utmost extensively used metric for assessing author productivity and scientific journals quality. The *h*-Index, created by physicist J.E. Hirsch (2005), is a measure of both quantity (number of documents) and quality (number of citations). The *h*-index value of a document indicates that the article has 'x' number of documents that have been cited at minimum 'y' times.

2.2 Learning Analytics

According to Siemens & Gasevic (2012), educational data analytics is the assessment, gathering, reporting and investigation of data about students and their perspectives with the goal of greater perception and enhancing knowledge as well as its environments. Learning analytics is a new area that focuses on using advanced analytic tools to enhance the educational process (Elias, 2011). It applies to a variety of other scientific fields, including educational data mining, academic analytics, action analytics, web analytics, business intelligence, and is closely linked to them. Moreover, learning analytics is one of the most critical developments in technological advancements in learning and teaching, stated in the NMC Horizon Report 2013 (Johnson et al., 2013).

Learning analytics concentrates on the learning experience in general (Siemens & Long, 2011). The invention, application, and adoption of new tools and techniques to enhance the pedagogical performance of the various teachers and students are all part of learning analytics research and innovations. Thus, as a result, it is not surprising that learning analytics has been existing in numerous topics of scientific articles. In addition, psychology, learning sciences, philosophy, cognitive science, mathematics, education, sociology, artificial intelligence, data mining, machine learning and others all have an impact on the area of learning analytics. Both

KMK3104 FINAL YEAR PROJECT 2

education and computer science are the most influential disciplines among key researchers in the related field (Dawson et al., 2014).

Based on Campbell, Deblois & Oblinger (2007) studies, they did advise that effective approaches are a major concern in learning analytics which there is a need to use the information obtained from the data to make approaches to enhance learning and also to produce 'actionable knowledge. This is discussed in the context of the learning analytics process. Furthermore, capture, report, predict, acting and refine are the five phases that have been laid out by Campbell & Oblinger (2007): -

- Capture: Data is generated and acquired in real-time from a wide range of sources and blended with student information. For example, forums, online chat or rooms, libraries, and other sources.
- Report: The data gathered is often used to create precise concepts for defining and assessing the progress of students. For a clearer vision, visualization is frequently used in learning analytics dashboards.
- Predict: The data is used to find predictors of student achievement and results, as well as to recognize students who are at risk. It is also used to make decisions about programmes and strategic planning, which are then used through institution choices.
- Acting: The data analysis method yields information that is used to develop effective approaches, such as tutoring or helping pupils who are going to drop out.
- Refine: The data collected is used in a cyclical method to develop the teaching and learning model that is currently used nowadays.

2.3 Learning Analytics Theoretical Model

With the advent of "big data," a theoretical model for placing educational data analytics within the high-level education was born. The global development of computer availability and instructional media, according to Mayer (2009), expanded the chances to improve learning processes. Learning analytics uses approaches from statistics, psychology, data mining, sociology, machine learning, and information science to analyse data acquired during education management, knowledge, and facilities according to MacNeill et al. (2014).

Shum and Ferguson (2012) developed the notion of common educational data analytics to discover patterns and behaviours at both the individual and group level. Educational data analytics combines data from both proper and semiformal educational settings (MacNeill et al., 2014). To help institutions implement learning analytics, Shum and Ferguson (2012) established the notions of macro, micro, and meso. Information distribution among institutes for a variety of objectives, together with benchmarking, is made possible through macro-level analytics. Meso-level analytics, which comprises business intellect analytics, are applied at the stage of specific universities. Individual learners can use micro-level analytics to track and evaluate process-level data (Shum & Ferguson, 2012).

2.4 Learning Analytics for Educational Purposes

Nowadays, the competition between the universities level growths. According to Van Barneveld et al. (2012), they stated that to ensure institutional performance at all levels, these universities must improve strategic and financial quality, extend nationally and international reach, develop new partnerships, and pay attention to requirements for greater transparency. Moreover, all institutions also must contend with the quickly evolving technologies that have emerged as a result of the technology generation's arrival.

In this case, more universities are employing learning analytics approaches to acquire information about the students' academic achievement, forecast future actions, and identify problem areas early on. Besides that, in the educational field itself, learning analytics is an appropriate instrument for reflecting students' learning behaviour and providing appropriate assistance from teachers or tutors. Thus, this support gives alternative methods of teaching as well as a method to focus on the student's learning behaviour.

In addition, some institutions have already implemented LA in a variety of courses to help students study more effectively. For example, the Alabama's University enhanced student preservation by developing a prediction model for vulnerable students according to a big dataset of statistical data. Purdue University also used predictive modelling depending on information from its course administration system to detect and intervene with students who were in danger. Moreover, the University of Northern Arizona developed a prediction model to estimate which learners might profit by which materials by linking resource consumption, the possibility of loss, and student achievement (Campbell et al., 2007).

2.5 Previous Studies

There are a few bibliometric studies on learning analytics in the literature. Waheed et al. (2018) and Philipps and Ozogul (2020) studies concentrated on analysis of citations and research network recognition. The research of Azevedo and Azevedo (2021) concentrated on a thorough review that took into account the ISI Web of Knowledge Core Collection, which is conveniently accessible and relevant to their research. According to Adeniji (2019) study, it presents a more in-depth analysis and is based on a Scopus database study. Adeniji study's approach divides the data into three periods, based on a thematic analysis of each time and the identification of different themes for each period. However, some of these past studies did not use the Scopus database. Adeniji did use the Scopus database but it is not updated until 2020.

CHAPTER 3

METHODOLOGY

This study's procedure is based on how collecting data and cleansed everything the refers to the final information collection, which is solid and prepared on being reviewed. The goal of this study is to focus on all works related to Learning Analytics (LA) that are available in the Scopus database, hence the subject and significance of the study must first be specified. Scopus was chosen only because of its reputation as the "largest individual abstraction and indexed library being developed" (Burnham, 2006) and the most complete citation and abstraction record of searches available (Ahmi et al., 2019).

The publications for this research were chosen based on the research process outlined in Figure 1. As of December 1, 2021, the information was taken from the Scopus database. To find all of the target papers, the following keyword combinations were used: "Learning Analytic" or "Learning Analytics" or "Educational Analytic" or "Educational Analytics" or "Educational Data Analytic" or "Educational Data Analytics".