Learning Analytics in Higher Education Development: A Roadmap

Olugbenga Adejo* Thomas Connolly
School of Engineering and Computing, University of the West of Scotland, United Kingdom
* E-mail of the corresponding author: olugbenga.adejo@uws.ac.uk

Abstract
The increase in education data and advance in technology are bringing about enhanced teaching and learning methodology. The emerging field of Learning Analytics (LA) continues to seek ways to improve the different methods of gathering, analysing, managing and presenting learners’ data with the sole aim of using it to improve the student learning experience and the study environment. In this paper, we try to explore the concept and salient features of LA potential in higher education and suggest strategies on how this emerging field can make use of data mining techniques alongside learners’ data to produce useful and informed decision making. Using the Technology-Organisation–Human frameworks, the paper investigates the roadmap for successful implementation of LA in higher Educational Institutions.

Keywords: Learning Analytics, Learning Management System, Higher education, Roadmap

1. Introduction

The world generally is experiencing a data revolution. There is an upsurge in the amount of data that is being generated from different sources, through different channels, in different varieties and at a great speed. The rate of growth of data is expected to double every month between now and 2020 (Khan et al, 2014). The higher educational sector has also not been left out of this “data deluge” era, there are great influx of data as well as emergence and adoption of new technologies. However, the major problem in the sector has been the inefficient use of this data to improve value to meet the educational market need as compared to other sectors such as marketing, banking, health, security and sport.

In addition to the students demographic and activities data, the widespread introductions and proliferation of online learning, intelligent tutoring and Learning Management System (LMS) have led to an unprecedented increase in data generation, personal data collection, learning or system data and interaction data. Beyond these are also a vast amount of data being collected and stored in educational blogs and forums, which are increasing daily and require special analytical tools for analysis and interpretation. However, the exploitation and useful application of these automatically collected users data for education development and learners benefits are still very limited (Greller et al., 2012; Dawson and Macfadyen., 2012).

Although, all these data have hidden values behind them and if well analysed can give deep insights into the educational setting and can reveal what teaching methods and academic interventions are most likely to enhance learner performance and retention (Wook et al., 2009). This simply means that the new set of the problem created by the “Big” educational data era needs a new set of solutions. One of such solutions designed to help the higher education system is Learning analytics.

Learning analytic (LA) is an important and emerging area of Education data Mining (EDM) and Technology Enhanced Learning (TEL) that comprises of different academic disciplines such as education, psychology and computer science (Papamitsiou and Economides, 2014). LA officially emerged in 2010 as a new domain of education and computer science and makes use of existing techniques such as Artificial Intelligence, business intelligence, statistics, mathematics and visualisation (Ferguson, 2012). LA specifically came out of EDM to address the issues of learners’ progression and attrition using the learner’s data. Ferguson (2012) and Clow (2013) in their earlier studies compiled a comprehensive summary of definitions of LA. However, one of the most commonly cited definitions was the one given by Learning Analytics and Knowledge (LAK, 2011) and asserted by The Society for Learning Analytics Research (SoLAR) in which LA was defined as “… the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimizing learning and the environments in which it occurs”. From this definition, it can be deduced that LA has no data of its own but exists to handle and analyse large educational data sets for learners’ benefits.

EDUCAUSE as cited by Bichsel (2012, pp.) subsequently defined LA as “the use of data, statistical analysis, and explanatory and predictive models to gain insights and act on complex issues” relating to learners. These two definitions, which are now generally accepted for learning analytics, are different from EDM, which the International Educational Data Mining Society defines as “an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in.”

However, Del Blanco et al. (2013) describe LA differently as “a discipline that gathered and analysed
educational data with different purposes such as seeking a pattern in the learning process and trends or problem in student performance”. The authors pointed out that LA collects huge amounts of data from student actions, courses and learning tools. Their actions include their interaction with other learners, instructors and learning content while the learning tools range from traditional LMSs to MOOCs (Massively Open Online Courses) such as Udacity and Cousera. LA makes use of different data mining techniques, statistics and visual analytics techniques to gather useful information From the various definitions of LA complied by Ferguson (2012) and Clow (2013), the words common to all are “huge”, “vast”, “voluminous”, and “large” datasets, which are what big educational data is all about- large volumes of data. Hence, it can be deduced that LA is in existence solely to handle the issues of big learner data and how learners can benefit from it. However, in this thesis, LA will be viewed as the collection, storage, and analysis of data from learning systems (LMSs-Moodle) in order to gain useful and intelligent insights for decision making that will have a lasting impact on learners and the HEI. In other words, the main intention of LA in this research is for gathering learners’ data for the development of models, algorithms, and processes that can be widely used and generalised to enhance learners’ performance.

Many researchers have investigated the use of LA for student benefits; Campbell and Oblinger (2007) relate LA to making an informed decision about student progress and retention and noted that “Retention of students saves institutions the cost of recruiting students to replace those who withdraw without completing a degree”, MacFadyen and Dawson (2010) investigate its use in predicting student performance using learning management system data while Jose et al., 2016 research into it use for progression analysis of students in HEIs. These studies assertions were corroborated by Dietz-Uhler and Hurn (2013) who explain the importance of LA to include assisting educational institutions in increasing student retention, improving student success and progress encourage personalised learning and also ease the burden of educational accountability.

However, Wise, Zhao and Hausknecht, (2014) identified two approaches to Learning Analytics which are of great importance in the use of it, these are;

1) Embedded analytics, which is the integration of traces of activity into learning environment itself which can be used real-time in guiding learner participation. This can be used in metacognitive monitoring during participation in learning activity.

2) Extracted analytics, which is the integration of traces of activity in a learning environment and represented in a separate format or platform; e.g. a dashboard.

Ferguson (2012) also identified three different groups in LA with overlapping interests to include governments, educational institutions and teachers/learners and later described that each group has direct and indirect impacts on the learners and their academic success.

The groups and their impact are shown in Figure 1

In a similar study, Greller et al 2012 proposed a framework of LA with benefit/information flow among the major stakeholders, the institution (administrator), teachers and students.

The typical uses of LA to higher education are

• For learners career development and progress measurement (Abukhousie 2014)
Learning adaptation and personalization to suit the need of each student (Arnold et al., 2012)
Educational games development (Miroslav, 2013 and Chaudy et al., 2014)
Development of Course signal
A convenient tool for assessment of teaching and learning in students
Similarly Larusson and White (2014) enumerated the uses of LA to include
Enhancement of student and institutional performance
Assist in assessing and giving attention to student at risk
Help instructor to assess and develop their own strength
Help institution to make efficient use of their learning resources.

The general benefits of the application of LA to the different stakeholders are summarised in Table 1.0
In addition, the tools available for LA are limited in their capability. Atif et al., 2013 summarises the
tools and approaches used in LA in higher education in their work and also compared the capabilities of four
different tools developed and used at Australian Universities. It was found out that some of the tools still have
limited capabilities in achieving their goals. Almosallam et al., 2013 provided an overview of these tools and
applications from LAK11 to LAK 13 Conferences and some of the tools mentioned include E2Coach, Learn –B,
Course Signal, SNAPP, café, AAT etc

However Dawson et al., 2014 noted that despite the wide-open and complex landscape of LA with the
consequent benefits, much focus of LA is still on univariate analysis which can be described as “low hanging
fruit” of LA potential compare to deep, multivariate analytical work with much deeper meaning that can be
done with LA. This collaborated the previous statement reported by McKinsey about the lack of extensive
strategic application of LA in the higher education sector as such there is a need for defining a strategic roadmap.

Table 1: The overview of LA multidimensional benefits framework

<table>
<thead>
<tr>
<th>Benefit dimension</th>
<th>Description</th>
<th>Sub- benefit impact / implication</th>
</tr>
</thead>
<tbody>
<tr>
<td>Governmental benefits</td>
<td>Impact of LA technologies use by teacher on government and society</td>
<td>• Reduce unemployment and other vices&lt;br&gt;• Improve rating for the government and state.&lt;br&gt;• High return on investment.&lt;br&gt;• Long term benefit of increase in tax revenue&lt;br&gt;• Long-term increase in GNP as a result of increase income of graduates&lt;br&gt;• Lower government expenditure on social welfare.</td>
</tr>
<tr>
<td>Institutional/organisational</td>
<td>The benefits that accrue to the institutions when thy make use of LA tools and technologies</td>
<td>• Changing in work patterns&lt;br&gt;• Facilitating improve educational/ learning delivery&lt;br&gt;• Improve ability to obtain funding and increase tuition revenue for the Institution as retention increases.&lt;br&gt;• Improve institution decision making&lt;br&gt;• Improve retention and graduation rate&lt;br&gt;• Improve School reputation and rating&lt;br&gt;• Reduce cost of re-recruitment for replacement.&lt;br&gt;• High return on investment.</td>
</tr>
<tr>
<td>benefits</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operational benefits</td>
<td>This is the benefits obtained when improved method of learning approach and operation are adopted</td>
<td>• Easy of analysis of student data&lt;br&gt;• Ease of monitoring students using LA dashboard.&lt;br&gt;• Help in suggesting an early intervention to a student at risk.&lt;br&gt;• Improve workforce productivity&lt;br&gt;• Students get the necessary assistance on time.&lt;br&gt;• Help student progression.&lt;br&gt;• Reduced debt as a student will acquire the means (the degree certificate) to pay it.&lt;br&gt;• High return on investment</td>
</tr>
<tr>
<td>Learners benefits</td>
<td>How does the student gain from the use of LA by the teachers</td>
<td></td>
</tr>
</tbody>
</table>

2. What is different about Learning Analytics
Educational administrators and technologists, especially those in higher education institutions, are still sceptical
about the major differences in LA as compare to other fields that makes use of educational data such as
Education Data Mining (EDM) and Academic Analytics (AA), after all they are all focused toward the success of students and the education sector. EDM has been defined by the International Educational Data Mining Society as “an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in.” Siemen et al. (2010) distinguish between EDM and LA, in areas such as their origin, stakeholders, adaptations and personalization, reduction to component and holism, which deals with understanding the system as a whole and its complexity. Piety et al. (2014) distinguish between Academic Analytics, Learning Analytics, Educational Data Mining, Personalization, and Systemic Instructional Improvement and pointed out that the major overlapping area of all these disciplines is in the engagement of researchers on how to exploit “big educational data” to improve education. Other areas of similarity are found in the fact that they are all deeply rooted in digital learning environments and online storage of data.

However, it should be noted that the differences between these fields as enumerated below are largely organic and actually developed from the views, interest and values of specific researcher rather than reflecting the sharp division or antagonism views. Baker et al., 2012; Almosallam et al., 2013 and Linan et al., 2015 gave comprehensive differences and similarities between EDM and LA. These are summarised below.

Table 2. Tabular Representation of the differences among the three analyses discipline

<table>
<thead>
<tr>
<th>Education Data mining</th>
<th>Academic Analytics</th>
<th>Learning Analytics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Reduced focus – Teachers and students</td>
<td>Reduced focus – Educational institutions</td>
<td>Holistic focus- Teachers, Learners, educational institutions.</td>
</tr>
<tr>
<td>2) Automated adaptation and method of interpretation</td>
<td>It involves automatic iterative process</td>
<td>Support human intervention and interpretation human data</td>
</tr>
<tr>
<td>3) Rooted in software and student modelling.</td>
<td>Focus on administrative concerns that affect higher education.</td>
<td>Focus on systemic intervention</td>
</tr>
<tr>
<td>4) Focus on learning as a research</td>
<td>Focus on enrolment management &amp; decision making</td>
<td>Focus on the aspect of education beyond learning</td>
</tr>
<tr>
<td>5) Make use of textual data for analysis</td>
<td>Make use of textual data for analysis</td>
<td>Textual analysis of data is rarely seen</td>
</tr>
<tr>
<td>6) Focus more on techniques and methodologies</td>
<td>Focus more on techniques and methodologies</td>
<td>Generally, focus on the application of analysis.</td>
</tr>
<tr>
<td>7) Make use of Data mining techniques</td>
<td>Make use of statistical techniques &amp; predictive modelling</td>
<td>Make use of Quantitative methods, data mining techniques and visualisation tools</td>
</tr>
</tbody>
</table>

3. The Unresolved Challenges of Learning Analytics in Higher Education

Learning analytics are designed to be of immense benefit to the organisation, teachers and learners in the areas of learning personalization, performance enhancement and feedback, student empowerment, learning methods design and consequently improved educational decision making. However there are diverse challenges mitigating against the application and implementation of it and if these challenges are not adequately addressed now, they will be a great impediment to the development of the sector and meeting LA expectation. The study tries to answer these following questions –

- What are the major challenges in implementing LA technology in HEIs?
- What are the roadmaps that need to be followed to achieve success in LA implementation?

These challenges identified by the study therefore are;

a) Data and database heterogeneity – Eighty percent of educational data are unstructured, in a different format and are spread around multiple, heterogeneous databases. These huge datasets are most time difficult to integrating and use hence needs to find a different approach for this multi-database mining and analysis. This will continue to be a major future challenge of LA as learner data is growing at a geometrical rate.

b) Technology architecture impact on data access – Presently, the field of LA lack the necessary technology architecture development framework as it relates to technology specifications, models design and other infrastructure architecture. This poses a great challenge to the development and use of LA technology.

c) Data ownership – The data collected from different databases of HEIs and those collected through survey questionnaire do have issues of ownership. Who owns or has the legal right to the data that are being used for analysis? How can the data be used, stored and for what duration? These and other
questions need to be clarified with regard to learners data, if not it will act as a major impediment in accessing and analysis of data in future.

d) Privacy issues- As a result of huge data set involve in learning analytic and access to these data by different categories of people vis a viz administrator, teachers and other HEI staff, there is need to address the challenges of security, privacy, Access and data profiling. The educational administrator must, therefore, establish a guideline to monitor access and use of learners data.

e) Consensus Research framework – Presently there is no standardised framework on LA application in educational system

f) Generalisation of application and tools - Many HEIs have already started implementing and utilised the potential LA in their system. Among them are those that are developed LA tools such as Purdue University (SIGNAL), , University of Wollongong (SNAPP), University of New England (UNE)- (AWE), Open University Australia (OUA) - (PASS) etc However most of these technologies are specific to the university and are not easily adaptable to others due to different social, economic and geographical factors affecting them. This is a major challenge that needs to be addressed for LA to be successful

g) Inadequate training of staff on the use of learning analytic tools.

4. The Roadmaps

Many questions have arisen relating to the future impact, direction and how to make LA achieve its ultimate goals. Researchers have shown that the impact, value, processes and technology involved in LA is not well understood and as such the potential and beneficial application of LA in the HEIs are undermined (Greller and Drachsler., 2012). Even though the research area is just emerging, it is facing its own challenges as any other research areas. In fact, one of the major theme for the conference of Learning analytic and Knowledge for 2014 is “Learning Analytics at crossroads” as such the new research area require a fairly structured and sequential process of development that meet the required educational needs. Therefore the first major approach is for the various stakeholders in the educational sector to identify and understand the strategic and potential value of learning analytics, rather than just focusing on technological implementation of it alone. Hence the paper tries to explore and proffer the different roadmaps that will lead to the success of the new research area. In his paper, Garcia et al 1997 presented standardised structured methods for designing and developing roadmaps, these are grouped into three parts- preliminary planning phase, development of the Roadmap, and follow-up activities. However, for this research, I have included other two phases. This is presented in Figure 2 below

Phase 1: Preliminary phase - This preliminary phase consists of identification of the present challenges/gaps and conditions that needed to be met to successfully implement LA in HEIs. It is the planning stage that involves the development of strategies for knowledge acquisition and accumulation about LA, its data sources and database available. It also involves the identification of major stakeholders and necessary technologies/infrastructure to address the gap, identification and assessment of the impact of action/inaction of each stakeholder on the success and development of LA as well as development of theoretical working analysis of LA. However, it should be noted that the first major step in this phase is the formulation of the major educational goals to be achieved with LA, how to achieve it (execution) and the performance measurement. While the second is making that critical decision on the dataset to be incorporated in the analytic as well as data integration from heterogeneous databases. Generally the preliminary stage deals mostly with planning.

Phase 2: Security Privacy and Compliance Phase - In this phase, adequate identification and understanding of security, privacy and legality issues relating to the accessibility, use and storage of learners data are focused on. Benchmark, guidelines, policies and standard for privacy, data protection, security, ethical use of data and other legal compliance are created.

Phase 3: Roadmap development phase- This phase involves designing and development of the key strategies to meet the specific requirement of LA development. Three major factors or roadmap strategies have been identified as necessary to ensure successful implementation and development of LA. However, these strategies must be integrated together in order to achieve the goal. These strategies are going to be examined under three sub-heading:

a) The technological factors
b) The organisational factors
c) The human-environmental factors

a) The technological roadmap of LA in higher education - Technology which is the driving force of analytics and organisational change as such its role in LA development cannot be overemphasised. Presently technologies application in LA is in form of “Black box” solution that is of negligible use to HEIs. In addition capturing multi and heterogeneous data of learners from different databases with the existing technology is challenging.

However, it should be noted that each HEIs have a different requirement based on location, the purpose
of establishment, mode of lecture delivery etc. Hence, the development of an appropriate template for implementation of LA. The LA technologies development road will, therefore, involves the technology need assessment and technology Development Plan, where the available technology are assessed against the needs, requirements and goals of the educational institution before selection. Also, there is need to identify alternative technologies that might of great use in meeting the needs. As such the following should be done by the institutions:

- Defining the requirement of the institution
- Literature review and understanding of each Learning analytic tools and technologies
- Evaluate and understand features and attributes as well as capabilities/limitation of each LA tools over the next 5-10years
- Understand the dynamics of the technologies for future use
- Aligning the attributes with the requirement and goals of the organisation.
- Adoption and designing of the target technologies
- Assess the performance of the technology in order to realign technology if necessary

b) The Organisational roadmap of LA in Higher Education- Organisational dimension in the success of LA technologies in HEIs considers the relationship that exists between various components of the organisation and the implementation of the technologies. These organisational components include organisational policy, operations, management structure and culture as well as the institution strategy on technology implementation. Success in the use of LA tools requires the development of comprehensive strategic and viable education plan. The study proposes four major organisational elements that should be considered as a roadmap for successful implementation of LA. This is represented pictorial below

![Organisational factors for implementing LA](image)

**Figure 3: Graphical representation of organisational factors for LA implementation**

Hence the following steps should be followed

- Guide process for change implementation
- Review LA infrastructure and fill any existing gap
- Assess compatibility with any other technologies presently in use for learning
- Establish institution training program
- Create a conducive environment for the implementation of LA tools.
- Ensure adequate support from the educational management and administrator
- Establish standardised control and monitoring of use LA technologies and data.

c) The Human-Environmental roadmap of LA in Higher Education

The major key to effective utilisation, implementation and interpretation of LA technologies to achieve organisation goals is to equip the key personnel to use learning analytics. This should have a top-down approach where the management is committed to the successful implementation. Similarly, the key personnel should have the relevant professional skill and competencies, such as the ability to analyse, think and make appropriate interpretation of the result obtained from LA technologies. Because the incorrect interpretation of result report can lead to the wrong judgement which may give false hope or cause emotional trauma to the learner. Therefore it is very important for educational administrators, counsellors and teachers to be provided with basic analytical training in areas such as data mining, business intelligence, graph analysis/interpretation and statistics. These should be cascaded down through the educational system

**Phase 4: Adoption and evaluation Phase**

Once the strategies and LA methodological system have been fully implemented in the system. It needs to be evaluated to assess the strengths and weaknesses of the tools as well as the impact on the legacy technology, organisation and learners. The assessment carries out on the technologies include Functionality, Usability,
performance and supportability tests while both the positive and negative impact on the organisation are measured. Also, the risk, barrier and limitation to implementation are assessed.

**Phase 5: Monitoring and control Phase** - During the implementation of different stages of the roadmap, the technology progress should be monitored and performance assessed by measuring the performance of the tools against organisational performance benchmark, student performance, security strategies, the rate of retention and graduation. A system that will constantly monitor the performance and integrate application performance should be installed and implemented. There is also the possibility of review of phases before the final execution plan is developed.

**5. Conclusion and Recommendation**

In this paper, we have discussed the uses and impact of LA in a higher educational system with specific focus on the United Kingdom Higher education. The paper also provided a better understanding of how HEIs can make use of LA tools as a mean of transforming educational sector and learners performance in order to achieve the desired educational goals. Some of the benefits of LA have been enumerated to include governmental, organisational and learners benefits.

While the major impediments facing the success of LA implementation were discussed. It is expedient that HEIs in the country make necessary policy and resolution to enhance learner data privacy and security. Also, effort should be made to mitigate the privacy and security risk associated with learner data accessibility especially to different teachers and education administrators.

This theoretical review research found out that most HEIs in the UK are not using LA to its optimum level to assist the learner in their effort to succeed and therefore recommend the review and adoption of LA to HEIs. Furthermore, the paper recommends that there should be an established standard in the use of LA in HEIs. This will help to reduce the issues of the privacy breach and legal issues that might arise as a result of its use. In addition, the educational administrator should advocate policies that promote training of teachers on the use of LA tools.

Finally, even though it is acknowledged that the paper has the limitation of not providing empirical evidence, however, it has actually provided a framework and roadmap for the implementation of LA tools as well help educational administration to estimate the future direction of LA and their relevance to different stakeholders. In addition, it should be noted that the foundation to generate the necessary benefit in LA application is a link to the proper integration that exists among learner, teacher and the institution.

In future, it would be advantageous to develop a dashboard that is more comprehensive to help teachers and education administrators to have a quick snapshot of student at-risk of attrition so that appropriate help/intervention can be given. Further empirical research should also be carried out by measuring the rate adoption/acceptance of LA tools in the country as well as the development of more refined learning analytics implementation model.

**References**


Garcia, M.L. and Bray, O.H., 1997. Fundamentals of technology road mapping (pp. 3-31). Albuquerque, NM:
Sandia National Laboratories.