Learning Analytics in a Shared-Network Educational Environment: Ethical Issues and Countermeasures

Olugbenga Adejo*
School of Engineering and Computing
University of the West of Scotland
Paisley, United Kingdom

Thomas Connolly
School of Engineering and Computing
University of the West of Scotland
Paisley, United Kingdom

Abstract—The recent trend in the development of education across the globe is the use of the new Learning Analytics (LA) tools and technologies in teaching and learning. The potential benefits of LA notwithstanding, potential ethical issues have to be considered and addressed in order to avoid any legal issues that might arise from its use. As a result of this, Higher Education Institutions (HEIs) involved in the development of LA tools need to pay particular attention to every ethical challenges/constraint that might arise.

This paper aims to identify and discuss several ethical issues connected with the practice and use of LA tools and technologies in analysing and predicting the performance of students in a shared network environment of HEIs. The study discusses the four ethical issues of Information and Communication Technology namely Privacy, Accuracy, Property and Accessibility (PAPA's Model) as well as other approaches to explain these future concerns. The paper also presents the empirical evidence of the views of students on the analytical use and storage of their data.

The results indicate that even though students have high trust in the privacy and security of their data being used by their institutions, more than half of the students have ethical concerns with the accessibility and storage of their data beyond a certain period. In the light of this, generalised strategies on ethical issues of the use of learners' data in an HEI shared networked environment are proposed.

Keywords—Learning Analytics; Student's data; Emerging technologies; Ethical Issues; Higher Education

I. INTRODUCTION

Learning analytics is an emerging technology that is becoming of great use in various areas of education - from the pre-school level to the tertiary level education. Presently, some researchers have also identified its uses in healthcare education and in educational (serious) games [31, 22, 29, 6, and 23]. The adoption of LA technologies has been very dynamic and ever changing as the nature of ICT changes globally. Its importance and benefits are seen in all sections of education, from traditional face-to-face teaching to blended education and e-learning [2]. In the same empirical study, Ali et al. [2] investigated factors influencing adoption of LA tools and found that intuitive interfaces, students' learning interaction, perceived usefulness of the tools, sustainability of usefulness among others have great influence on adoption of the tools while the educator's pedagogical role has no influence. However, its implementation in the UK is still in the nascent stage, presently nearly half of the HEIs (47.2%) have not implemented LA while another 34% are just

considering implementing it. In total, only about 18.9% have either fully or partially implemented it [14]. In predictive analytic use of LA in measuring students' performance, over 95% of HEIs are not using it at all with a mere 2.5% making adequate use of it. This low implementation rate has been attributed to the level of understanding of the possible benefits of LA in HEIs apart from the problem of the high cost of initial implementation. In total, only 22.6% of management in HEIs understand the short and long terms benefits of LA use in their institutions [14].

Therefore, one of the recurrent questions from this emerging area of education is how to increase the motivation to use it effectively in HEIs. However, the bulk of the responsibility for its use is on the major stakeholders - the administrators (Institution), the teachers and the students [12, 3]. These three major stakeholders have great responsibilities in the development and use of LA technology.

- **The Administrators:** Identify the technology, propose its use, provide an enabling environment, propose the implementation strategy and monitor the use of the tools. Other responsibilities revolve around administrative decision-making.

- **The Teachers:** decide on the importance, acceptance and adaptability/personalization of the tool to meet the interest and preference of the learner. The teachers assess the suitability of the tools and give feedback on how to improve the functionality and performance.

- **The Students:** determine the impact of the use of the tools on their learning and learning environment. The students give the necessary feedback on the advantages of the tools on their performance.

Other stakeholders include the course developer /researchers, computer/network administrators, technicians, data analysts. Therefore, LA implementation is human–centred and thus requires human cooperation for its effective adoption. Though the use of LA is not the panacea to all student problems, it may form a useful tool for identifying students at risk of failing.

However, ethical issues need to be considered with implementing LA tools especially now that HEs have systems that link students' demographic data with their educational performance data as well as their interaction with online learning resources to predict their future performance or even recommend material for them [29]. This is often referred to as

longitudinal unique identification that helps connect data across platforms or systems [11]. HEs are a good example of such complex adaptive systems (CASs) where data, files, programs and other resources are collected and shared across different systems by different people [13, 16]. This is referred to as a shared networked environment. Hence, stakeholder involvement is necessary from the design stage to the implementation stage of LA tools with focus on unintended consequences and challenges that might arise with this emerging technology.

Learning Analytics, like any other emerging technology, poses new challenges emerging from the complexity of its uses, security, accessibility, privacy as well as other ethical issues. These ethical issues and other practical questions about Learning Analytics ethics are still unresolved [24, 7]. According to [24], the present legal framework in HE cannot handle properly the issues of ethics, security and privacy concerns that might arise within LA in the future. It will, therefore, be of great benefit to identify these challenges and find solutions to them at an early stage of development before they become more difficult to address later. Nevertheless, it should be noted that ethics in practice might be a bit difficult as there will always be a situation where the new framework may be found wanting in a real–world situation as technology progresses [10].

The unethical use of LA technologies by stakeholders (teachers, administrators, system administrators) as well as other system users will be increasingly invasive toward the lives of the learners and may further lead to serious problems such as exposure of confidential data to hackers, cyberbullies, or other criminals. In addition, the convergence of other disciplines (education, psychology, computing) in LA has made the coherent discourse on LA ethics difficult as well as making Mason's (1986) ethical model look outdated. Therefore, it is expedient to know the potential benefits and application of LA in HEIs but also to understand the potential ethical issues in HEIs and how to offer solutions to them.

Hence this study primarily investigates ethical issues associated with the implementation of Learning Analytical tools in HEIs. Three main research questions are proposed for the study;

- Q1- What are the main ethical concerns around the use of LA tools and technologies in HEIs.

- Q2- What are the views of learners on the ethical use of their data by HEIs?

- Q3 – What are the coping strategies to handle LA ethical issues in HEIs?

A good understanding of the LA ethical issues and the strategy to handle them will be of significant use to the various stakeholders in HEIs and the educational sector in general. The findings from the study might help to facilitate the development of a general policy framework for LA ethics in HE, facilitate training programs for students and other stakeholders as well as help in improving the design of LA tools and technologies to be more ethically complaint.

II. THEORETICAL BACKGROUND

A. Defining Learning Analytics, Learner Data and LA Tools

There have been several definitions of LA that have emerged, however, the definition given by the Society of Learning Analytic Research (SOLAR) at the 1st Conference on Learning Analytics and Knowledge has been widely accepted. They defined LA as *"the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising 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 derived meta-data and other large datasets of learners for their benefit.

Another definition was given by Del Blanco et al. [8] who described LA as *"a discipline that gathered and analyzed educational data with different purposes such as seeking a pattern in the learning process and trends or problems in student performance"* while Greller and Drachsler [12] defined it as an academic domain *"that focused on learners, their learning processes, and behaviors"*. From all these definitions it can be seen that LA focuses on learners and the huge data collected about them for meaningful pattern discovery.

Learner data or student data can be defined as known or assumed information or facts, an observation about the learner collected for the purposes of analysis and decision making. It can also be described as the extraction and analysis of both static and dynamic learner data from different heterogeneous systems (databases) for use in the improvement of the learning process and the learning environment. Static data is from the university record systems while the dynamic data is from the systems such as the learning management system, virtual learning environment or social media. Learner data is an integral part of a learner's life and gives a clear picture about them, their past, present and the future and as such it is of great value to them. Alber [1] noted that learner data can be collected using formative assessment (quizzes, debate, question and answer), observation, summative assessment (projects and exams) and previous standardised test scores as well as cumulative files. However, the sources of learner data available to HEIs is continually increasing and this can be classified into five major groups:

- Demographic data: This data includes age, gender, place of birth, location, education, employment status.

- Behavioural data: This includes student interaction data from discussion forums, online activities, blogs etc as well as video and audio records of groups or individual activities.

- Assessment data: This includes data from coursework, assessment and examination.

- Financial data: This is the financial information store about the learner and it includes the sources of funding, fee status, tuition update, support allowances.

- Historical progression data: This is the information store about students after graduation and includes

degree type and class, employment data, location, forwarding address/contact, continuing education data, and alumni.

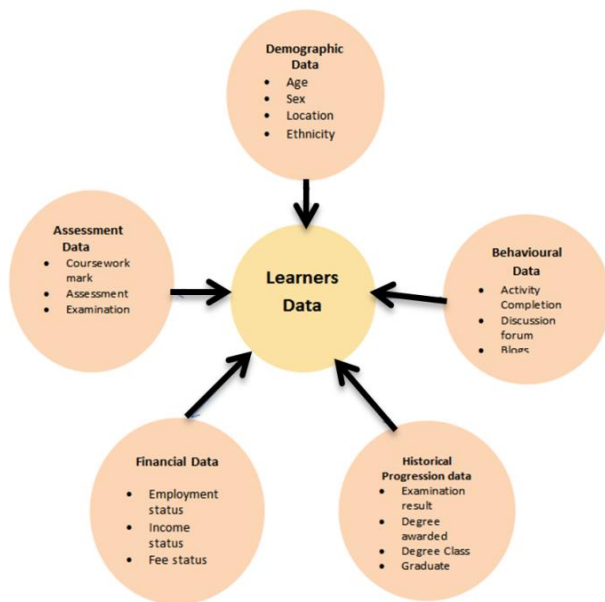The diagram of common sources of learner data is presented in Figure 1

Fig. 1.   Different sources of learner's data

While HEIs have always had access to data to work with, the increase in data being generated over time is unprecedented and associated technologies and analytic capabilities are trying to cope with the data increase [21]. In their study, Prinsloo et al [21] described LA as the "new black" box and learner data as the "new oil. They further explained that as there is an increase in learner data, there needs to be an understanding of the complexities of analysis of these data. This knowledge should facilitate the development of enabling technologies that will help gain an insight for the improvement of student performance, success and retention.

LA tools have been categorised variously by researchers. Siemens [24] grouped LA tools into two broad classes: **Commercial tools** developed by companies such as SPSS, NVivo and Stata and **Open Tools** such as Rapidminer, R, SNAPP and WEKA. Slater et al [26] also reviewed forty EDM/LA tools and categorised them into Data manipulation tools (MS Excel, EDM workbench, Structured Query Language), Algorithms analytical tools (Rapidminer, WEKA, SPSS,KEEL, Orange), Visualization tools (D3.js. D3.js, Tableau) and specialized tools such as Text mining, Coh-Metrix, WMatrix, Linguistic Inquiry and Word Count, LightSIDE. Other analytic tools that are more educational or system-specific include: Ellucian, Desire2Learn, Purdue University Signal for early warning about student performance, University of Maryland's "Check my Activity" tools for comparing learning activity of learners in LMS and Edith Cowan University's early warning tools called Connect for Success (C4S). A summary of all these tools was presented in a special study of LA tools and approaches by [4].

These LA tools used along with the learner data make analysis and assessment of students possible early in the semester without waiting for the tests or examination results at the end of the semester/year; ie. all the valuable information about students can be made available at the beginning or during the course of study so it can be used to assist their learning and learning environment. However, it should be noted that LA has no specific tools of its own but makes use of data mining tools for analysis and interpretation, which make it of great necessity to develop LA specific tools.

III.   THE MAJOR ETHICAL ISSUES IN THE IMPLEMENTATION OF LA IN HEIS

Learning analytics is a recent trend in education data mining (EDM) that makes use of learners' data for learners' benefits. In HEIs, the benefits of its implementation are quite clear and such benefits have significant advantages for all the stakeholders. However, despite the laudable benefits of the implementation of LA tools in HEIs, there are still some ethical issues that might arise that need to be properly addressed.

Ethics has been defined by Bynum [5] as various concerns for the protection and advancement of central human values such as life, health, security, happiness, freedom, knowledge, resources, power and opportunity. It is the fundamental decision on what is 'good,' 'bad,' 'right,' 'wrong,' 'ought', before applying actions [27]. As such, it is not just good to understand the benefits of any new technologies but the ethical challenges that may arise from it. Though the issues of ethics in IT have been developed since the 1980s, its importance and application to the new age and emerging technologies needs to be revisited especially in this era of "Big data" in education. LA ethical issues are concerns on how the learners' data that are available to the users can have an impact on learners' lives and uphold their moral values. In this regard, the following question arises: "has the use of the learner's data benefitted the HEIs and makes the learner happy or otherwise? Therefore, to answer this question correctly, all potential ethical issues should be identified and addressed during the course of the design, development and implementation of the technology.

Willis, Campbell and Pistilli [30] in their article discussed these ethical issues using the Potter Box Ethical model to answer various questions that an administrator must address when using learner data. The Potter Box ethical Model/Concept is an ethical decision-making tool made up of four quadrants (steps):

- Provision of empirical definition (ie. defining the ethical dilemma using empirical information).

- Identification Values (involves prioritising the ethical issues).

- Identification and appeal to ethical value (involves identification and application of ethical principles or reasoning that will help in making an effective decision regarding the dilemma).

- Choosing loyalties (making a choice of where loyalties lie).

From this concept, different responsibilities and recommendations are made to the stakeholders that will enable them to make an effective ethical decision. However, this paper examines the ethical issues using Mason (1986) PAPA's Model and other approaches that describe the various vulnerabilities of people and organisation in the emerging ICT era. These are discussed below:

Property - This term applied not just to the concepts of software or technological ownership but also to Data ownership as well as the right to information and value of information. The data collected and stored in different databases of HEIs, the one collected through interaction with learning management system (via trackers, xAPI) and those collected through survey questionnaire do have issues of ownership. Who owns the data and the information derived from it and how can the information be used? Thinking of these issues brings about the following questions.

a) Who really owns the data? The student? The University? The System provider?

b) How can the data be used?

c) Who has access to learner data analysis?

d) What type of learner data is stored in that database?

e) How and for what duration can the data be stored?

f) Which geographical areas can the data be used?

g) Should students be able to see what the institution sees about them?

h) What happens to the data after students completion of their studies?

These and other related ethical questions need to be properly addressed. Though the issue of data ownership may sound legal, there is a moral ethical issue embedded within it as well. Presently the legal system has not been able to address the challenges of data ownership and copyright in the use of the learner's digital information [28].

Moreover, Property involves protection of intellectual property rights (IPR). The increase in data usage and new technologies have made protecting IPR challenging and this might have an impact on the ethical use of LA technologies. It also includes knowledge and skills possessed by the user of the technology. The lower the skills and knowledge, the lower efficiency and the safe use of the tools. Stronger IPR ensures that the tools and technologies (especially computer programs or algorithms) are not easily or cheaply reproducible and provide protection and security to the data. However, this might affect the rate of future development of such tools. In addition, students should know about the fundamental laws that protect them, their data, their works as well as the use of their data or materials by others.

Accuracy - This ethical issue mostly arises in LA implementation where there are complexity and multiplicity of databases to access for analysis. As learner data sources, complexity and variety increase, the issue of analysis and computation of accuracy may be adversely affected. Continual system input, processing, and system output, especially in a shared network and distributed databases using different LA tools, can often lead to missing data, errors in data as well as

data mix-up or misplacement. Moreover, analytic software complexity can lead to inaccurate data being generated and used. Another major source of inaccuracy is in predictor selector and modelling which sometimes can involve complex statistical analysis. Therefore the question is how to ensure validity through the minimization of incomplete and inaccurate data. It is, however, important to note that, there might be legal issues involved when inaccurate or incorrect information about learners are used or distributed among systems.

Privacy – Privacy in the context of LA is ensuring the confidentiality of learners' information. Privacy protection is a widely accepted method of data protection and LA may pose privacy concerns as more data is being analysed especially where commercial tools are being used. One aspect of privacy is the inappropriate or unauthorised access to personal information. Every access to a database service by a stakeholder (user) means there is exchange of some information between the system and user as well as with various systems and even with the service provider. In addition, many commercial tool providers are looking for ways of making a profit from not just the students' data but also from the teachers' and universities' data and this is done through reselling of data to third parties [18].

In HEIs the following set of people has direct access to learner data: management, administration, student affairs and student counsellors, faculty and the students themselves. Hence, there needs to be an understanding by the learner of the type and extent of information they are implicitly or explicitly providing others access to. In as much as it might be viewed that all these people are helping the student to succeed, to some learners, it might mean an invasion of privacy. Unrestricted, unauthorised and privileged access to student data can lead to a breach of privacy, the source of which can be complicated to trace. Privacy breach in HEIs can also be the consequence of faulty protocol, database invasion as well as abuse by malicious users. This may involve divulging important confidential information about health status, marital status, personal information, private email correspondence and other information to a third party. Privacy issues usually occur during LA implementation due to lack of user control mechanisms over the networked environment as well as unauthorised access to data.

Accessibility – Accessibility in LA deals with authority and right to obtain learner data and the type of information one should have access to. The issue of accessibility relates both to the technology and the learner from which the data is derived. Learners should be aware and have access to the data and LA operation as well as the results of such operations performed on their data but presently this is not so. This often leads to lack of trust on the part of the learner. However, the issue of accessibility should be viewed from the learner's angle. Learners have the legal right to have access to their data, access the accuracy and correctness as well as being able to correct it when there is any mistake in the data held about them.

Trust – Trust enhances the implementation of LA as well as the technology acceptance of LA. The issues of trust by the

learner in the use of their data is another major concern that might seriously affect LA as the data might be in the hands of the third parties. What about the situation of distrust in which the student may be unwilling to provide data or opt-out of the data being used for analysis? These and more questions need to be properly addressed. Trust is not just related to the tools or technologies but also to the users. Though the HEIs might give assurances about trust, the situation might arise where some elements within the shared networked system might mishandle and mismanage student data leading to security concerns and consequently, such trusts have to be periodically reviewed.

Security - There are various security issues that need to be considered during implementation and use of LA tools. A security breach or risk may be due to data location compromised, data loss, unauthorised access or unintended or inappropriate disclosures/exposure of data. In a networked environment, the challenges of breach of security are greater. This may be as a result of switching-related problems where, in most time, the confidentiality is breached.

IV. SURVEY SCENARIO OF LEARNER ETHICAL PERCEPTION ON THE USE OF THEIR DATA

A. Materials and Methods

The study carried out a survey questionnaire using Survey Monkey in September 2016. The questionnaire emails were sent to students' mailboxes and non-probability sampling techniques was used in which every student have equal chance to respond as well as being selected. A total of 209 students at the University of the West of Scotland responded to the survey over a period of two weeks. The aim of the survey is to measure student perceptions on the use of their data and other the ethical issues relating to it. 77 participants (36.8%) were male and 132 (63.2%) were female while the modal age of participant ranges from 18 to 25 years. The questionnaire consisted of two main parts, first part collected demographic information while the second part was made up of six LA ethical-related questions, namely informed consent, security, awareness, trust, data accuracy and accessibility to data. The questionnaires responses were then analysed using SPSS version 22.0.

B. Results and Discussion

The survey was conducted based on the six questions in Table 1 using some of the items mentioned in Mason's four ethical issues of the information age [17]. A total of 151 students in level 8 to level 10 provided valid answers to the questionnaire for the study. The detailed results are presented in Table 2.

TABLE. I. QUESTIONS ON ETHICAL PERCEPTION AND CONSIDERATION

Question
Q1 Are you aware of the ethics and data protection policy of the use of your data by your Institution?
Q2 Do you trust your tutors and the institution mechanism on the use and security of your personal information?
Q3 Do you ever think your personal information privacy might be breached or information used inappropriately by your institution?
Q4 Would you want your personal or analysed information to be kept for a longer period by the institution after your graduation?
Q5 Have you ever observed mistake or complained about the correctness of the personal information or data kept about you by the Institution?
Q6 Would you like to have access to analytical information stored about you by your Institution?

TABLE. II. PERCENTAGE OF RESPONDENTS TO EACH QUESTION

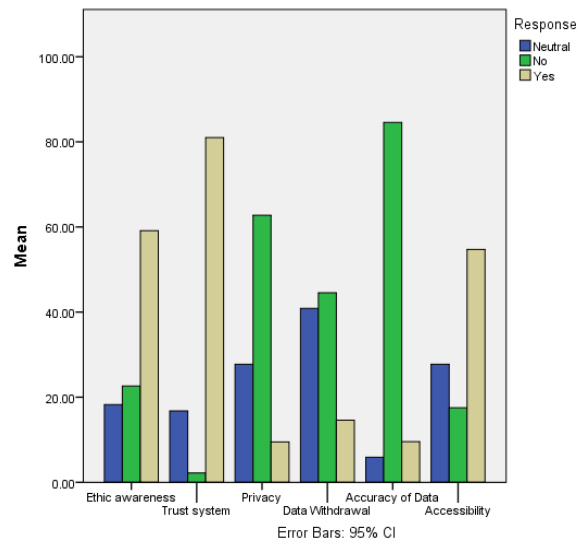| Scale (%)	Q1	Q2	Q3	Q4	Q5	Q6
No	22.00	2.67	62.67	46.67	85.23	18.00
Neutral	18.67	17.33	27.33	39.33	6.04	27.33
Yes	59.33	80.00	10.00	14.00	8.72	54.67

Fig. 2. Learners' Ethical Perception Results

Figure 2 shows that 33 students (more than 20%) are not aware of the ethics and data protection policy operating within the HEI and as such may not be aware of how their personal data might be used by the institution while another 28 students (18.5%) are not certain of whether such a policy actually exists or not. Though there is a sizeable number of students (90 students / 59.6%) that are aware of the data protection

policy but the general result shows that there is an inadequate awareness on the ethical issues and their rights by the students and as such might not know when there is an abuse of their rights or misuse of their information.

For the second question, majority of the students (121 students / 80.1%) trust the institution's mechanism for the security of their personal information and thus believe that it cannot be hacked or their information used inappropriately. For the third question, 94 students (62.3 %) think their personal information is protected while only 15 students (9.9 %) believe it might be breached. However, the opinion of students on question 4 shows that 70 students (more than 46%) want their data and analysed information to be removed from the institutional databases and never to be used for further analyses after their graduation. In addition, nearly equivalent percentage (39.3%) of students are undecided on this, signifying that a significant number of the students are not aware of either the positive or negative ethical consequences of the post-graduation data storage. This further points to the problem of awareness of ethical rights and privileges as discussed above in question 1. It can be deduced that over 80% of students (about 120 students) want their data removed after graduation strongly suggesting that HEIs should formulate a policy that will ensure periodic appraisal of consent on approval (withdrawal) of data use. The respondents' judgment on the issue of accuracy of data as depicted by question 5 shows that 127 students (84.1%) have never observed any mistake in their data while 10% of the students have observed inaccuracies. This shows significantly meticulous data collection and processing methods in the university. Question 6 on the issue of student data accessibility shows that 83 of the students (nearly 55%) still have some ethical perception regarding data accessibility and would like to have more access to their personal data and analysed information stored about them by the university. Again, universities should encourage and permit this.

The general consensus from this study is that there is a need for an up-to-date and complete reference on the issues of ethics in the light of the general growth rate of new LA technologies and increase in the volume of learners' data. Learner's awareness of the ethical issues relating to them should be enhanced through regular training or the use of game based IT ethics and security training.

V. THE COUNTER MEASURES

Since the use of LA technologies is still at a nascent stage of development, the proper use of learner data for analysis and prediction must be a matter of great attention. Personal student data should be protected in order to avoid loss of confidentiality, trust, and integrity by the students. However, the bulk of the responsibility of dealing with ethical issues in HEIs lies with administration and Government. The following are the proposed strategies for protecting learner data from the design of the technology to deployment of the technology.

Design - Tools design approach and methodology can play a major part in dealing with issues ethics, security and privacy especially at present when the LA discipline is just emerging and there exists a lack a standardisation [24, 8]. There should be ethical considerations guiding the design, development, and deployment of LA technologies in a shared environment. Smart LA technology is a major solution to the problem of lack of LA specific tools and potential ethical issues that might arise in LA implementation. However, it should be noted that some HEIs have designed their own specific technologies but there are still many ethical issues relating to their use. Future design of LA tools should consider the various aspects of ethics for education in a shared environment, which should include supporting technologies that address the problem of privacy, security, and trust. Below are some suggested areas where ethical considerations should be made during the design.

a) Tools should have a section where data can be made anonymous; this is known as value sensitive design (protecting what a person considered of value or sensitive to him or her).In addition, data can be encrypted to prevent unauthorised access.

b) Development of an efficient system that streamlines the techniques used for heterogeneous databases.

c) Tools should be designed to improve the interoperability, extensibility and reusability of data within the system without breaching privacy [9].

d) Tools should be designed in such a way that incorrect or inconsistency of data entered is flagged or highlighted even in different databases e.g. age, race, gender.

e) Tools should be designed to be able to adapt to increasing data sources.

f) Generally, LA tools and systems should be designed in consultation with students in order to address the ethical issues correctly

g) The tool should be designed to align with the institutional ethics standards.

Centralised control - In order to ensure the security of learners' data, there should be a well monitored and centralised control of access to student database system. However, the level of access should be layered to avoid a breach. The following control mechanisms should be carried out in networked environment

a) Use of Certificate of Authority – This involves registering and attributing authorities to all users, making sure all valid users are assigned global unique identities in before implementing access authorisation control.

b) Access authorization control.

c) Access Authentication.

This can always be achieved using single and multiple sign-in procedures.

Privacy mechanisms - In a networked environment, privacy mechanisms should be designed and developed using specific data mining and other algorithms. The privacy mechanisms should cater for the LA domain with focus on these two categories of people who might have access to the system – genuine staff and malicious workers (including hackers). To address the issues of privacy, policy on data usage in LA should be clearly defined. This should include the extent of access to learner data by staff, the use of sensitive data (religion, ethnicity), sharing of data with external bodies,

policy on anonymizing of data. Staffs who make use of a shared networked environment must understand their limitations, what they are allowed or forbidden to do. This should be well defined. In addition, to enhance privacy protection, the tool could prevent the execution of certain kinds of analyses where it would lead to the identification of individual students. In addition, staff who have no reason to use the LA tools should have no right to a system or databases where personal files and detail on students are held.

Reduced and restricted access - Reduced and restricted access should be enforced on all personnel that are not directly related to teaching and advising the student with their performance. Personal Identification Information (PII) can be implemented in order to get access to areas considered to be confidential and personal to students. There is a need to define a guiding principle of accessibility, who has access to a certain section of the network or database, and the kind of information that needs to be accessible to them. These questions can then be raised "which person or group of people should have more accessibility than others? And which criteria will be used in deciding the level of accessibility?" This should be properly defined for all users; needs-basis authorization should also be created while the proxy use of data is discouraged in all cases.

Informed consent and approval - To address ethical issues from the student's point of view, there should be transparency in the way learners' data are used. HEIs should always seek students' informed consent for collection and use of their data. This should include outsourcing the data to third parties, the right to opt out, withdraw from data collection, periodical renewal of consent to use data, consent to use data in a different way as well as protection of identity. The same process should also be applied to publicly available data about the students.

Policy framework development: Presently in the UK, only 10% of HEIs have adopted a code of practice or policy on the ethical use of data [14] and in order to avoid ethical breaches in the use of LA in a shared environment, new policies relating to IPR issues should be developed by HEIs. Just as many educational establishments have collaborated to develop standard e-learning content interoperability [8], similar collaboration should be geared towards establishing ethical standards for LA. The HEIs must put in place a well-defined access policy covering all stages of development. Without this, institutional change in the use of data may not come about. The policy should include guidelines and codes of conduct for staff using student data, including clear identification of who has access to learner data, the level of access, restriction. A good example of the policy is enumerated by Sclater in JISC's code of practice and rule of engagement for Learning Analytics. Developing a good policy framework in a shared environment will not only improve significantly the level of implementation of LA tools but also address various ethical issues that might act as a hindrance to its success.

In line with all these measures, institutions need to regularly organise conferences, workshop, and training for all academic, non-academic staff as well as students for proper understanding and knowledge about LA ethics and code of practice.

VI. CONCLUSIONS AND FURTHER WORK

The uptake and use of LA is expanding and as such the ethical issues related to LA practice need to be considered and addressed.

With the increase in the use and acceptance of LA technologies, the risks and consequences of e unethical use of learner data may increase especially in a shared networked environment. The current ethical policy framework used in most HEIs may not be sufficient to effectively address the potential issues of the use of these tools as well as the legal consequences. In addition to this, students (learners) are becoming more aware of their legal rights in relation to the use of their data and if a proper ethical framework is not put in place, there might be serious problems.

Therefore, this paper has examined the ethical issues arising from implementing of LA technologies in such an environment and proposed the adoption of enhanced practices and policy frameworks from the design, development to the point of deployment of LA tools that will ensure that the rights and responsibilities of learners respected. In addition, the study proposes a framework that will address the issues of accessibility, post-graduation informed consent and security of learners' data and has suggested the establishment of ethical rules to guide the users and necessary sanctions applied when breached.

One contribution of this study apart from the opening of theoretical discussions on ethical issues of LA in a shared network environment of HEIs is the establishment of the learners' views on accessibility to their personal data and analytical information. The research has shown that most learners would like access to the data kept about them by the HEIs as well as have the right to withdraw consent for their data use after graduation. However, the study has a limitation in that it was performed at only one institution in Scotland and with a small dataset, therefore to be able to generalise the result of the study a larger sample should be used and the study should cover a representative sample of UK HEIs.

Future work needs to empirically address other open challenges of LA ethics (such as benefits and dangers of giving data ownership rights to any of the stakeholders) as well as develop a framework and a system that will integrate and support students' accessibility to LA tools and their data. Also, more empirical work should be performed on ethics of LA tools using a student-centric approach.

Finally, it must be noted that ethics is gradually becoming a reference discourse in Learning Analytics and as such a clearly articulated framework, policy and model should be developed for it.

REFERENCES

[1] R. Alber, Three ways student data can inform your teaching. [Web blog].Retrieved from http://www.edutopia.org/blog/using-student-data-inform-teaching-rebecca-alber. 2011

[2] L. Ali, M. Asadi, D. Gašević, J. Jovanović, and M. Hatala, Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education*, *62*, pp.130-148, 2013

[3] E.A Almosallam,. and H. Ouertani,.Learning Analytics: definitions, applications and related fields. In *Proceedings of the First International Conference on Advanced Data and Information Engineering (DaEng-2013)*(pp. 721-730). Springer Singapore., 2014

[4] A. Atif., D. Richards, A. Bilgin, and M. Marrone, Learning analytics in higher education: a summary of tools and approaches. In *30th Australasian Society for Computers in Learning in Tertiary Education Conference, Sydney*. 2013.

[5] T. Bynum, Milestones in the history of information and computer ethics. *The handbook of information and computer ethics*, p.25, 2008.

[6] Y. Chaudy, T. Connolly, and T. Hainey, EngAGe: A link between Educational Games Developers and Educators. In Games and Virtual Worlds for Serious Applications (VS-GAMES), 6th International Conference on (pp. 1-7). IEEE, September 2014.

[7] M. Conde,. and A. Hernández-García, A promised land for educational decision-making?: present and future of learning analytics. In *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality* (pp. 239-243). ACM, November 2013.

[8] A. Del Blanco, A. Serrano, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, E-Learning standards and learning analytics. Can data collection be improved by using standard data models?. In *Global Engineering Education Conference (EDUCON), 2013 IEEE* pp.1255-1261. IEEE.

[9] A. Dyckhoff, D. Zielke, M. Bültmann, M. Chatti, and U.Schroeder, Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society*, Vol.*15*(3), pp.58-76, 2012

[10] S. Finlay, Predictive Analytics, Data Mining and Big Data Myths, Misconceptions and Methods Palgrave MacMillan publisher pp.89

[11] C. González-Sancho, and S. Vincent-Lancrin, Transforming education by using a new generation of information systems. *Policy Futures in Education*, p.1478210316649287, 2016

[12] W. Greller, and H. Drachsler, Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational technology & society*, *15*(3), pp.42-57, 2012.

[13] A. Gupta, and S. Anish, Insights from complexity theory: Understanding organizations'better. IIMB Management , 2009.

[14] HeLF UK HE Report on Learning Analytics 2015 (https://drive.google.com/file/d/0Bz7E74T5Am22bXpMc0hETDBadlk/view?usp=sharing)

[15] E. James, I. Willis, D. Matthew, Pistilli, and C. John, "Ethics, Big Data, and Analytics: A Model for Application," *EDUCAUSE Review Online*, 6 May 2013.

[16] B. MacLennan, Evolutionary psychology, complex systems, and social theory. Soundings: An Interdisciplinary Journal, Vol. 90(3/4), pp.169–189, 2007.

[17] R. Mason, Four ethical issues of the information age, MIS Q., Vol. 10(1), pp. 5–12, 1986.

[18] P. McRae, S. Varnhagen, and B. Arkison, The Alberta Teachers' Association. *Learning*, Vol *81*, p.01.

[19] N. Sclater, Code of practice for learning analytics, https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics, 2015

[20] A. Pardo, and G. Siemens,. Ethical and privacy principles for learning analytics. British Journal of Educational Technology, Vol. 45(3), pp.438-450, 2014

[21] P. Prinsloo, E. Archer, G. Barnes, Y. Chetty, and D. Van Zyl, Big (ger) data as better data in open distance learning. *The International Review of Research in Open and Distributed Learning*, *16*(1), 2015.

[22] W. Raghupathi, and V. Raghupathi, An overview of health analytics. *Journal of Health & Medical Informatics*, *2013*.

[23] A. Serrano-Laguna, and B. Fernández-Manjón, April. Applying learning analytics to simplify serious games deployment in the classroom. In *2014 IEEE Global Engineering Education Conference (EDUCON)* (pp. 872-877). IEEE, 2014,

[24] G Siemens, Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, p.0002764213498851, 2013.

[25] G. Siemens, and P. Long, Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE review*, Vol.*46*(5), p.30,2011

[26] S. Slater, S. Joksimović, V. Kovanovic, R.S. Baker, and D. Gasevic,. Tools for Educational Data Mining A Review. *Journal of Educational and Behavioral Statistics*, p.1076998616666808, 2016

[27] A. Thomson, and D. Schmoldt, . Ethics in computer software design and development. *Computers and Electronics in Agriculture*, Vol.*30*(1), pp.85-102,2011

[28] J.C .Thomson, K. M. Jones, and K. Arnold. Questions of Data Ownership on Campus. EDUCAUSE review, 2014

[29] C. Vaitsis, G. Nilsson and N. Zary, Visual analytics in healthcare education: exploring novel ways to analyze and represent big data in undergraduate medical education. PeerJ, 2, p.e683,2014

[30] J. Willis, J. Campbell, and M. Pistilli, Ethics, big data, and analytics: A model for application. EDUCAUSE review online. 2013. Retrieved from http://www.educause.edu/ero/article/ethics-big-dataand-analytics-model-application.

[31] H. Zhang, P. Zhang and R. Gilbert, Role of Machine Learning Analytic Tools for Survival Prediction of GBM Patients Using Clinical Parameters. H.H. Zhang. Proceedings of the American Society for Radiation *Oncology 57th Annual Meeting, Vol 93(3) pp. S142–S143,2015*