Towards a data archiving solution for learning analytics

Sarah Taylor
RMIT University
Melbourne
Australia
sarah.taylor@rmit.edu.au

Pablo Munguia
RMIT University
Melbourne
Australia
pablo.manguia@rmit.edu.au

ABSTRACT

Data solutions in the teaching and learning space are in need of pro-active innovations in data management, to ensure that systems for learning analytics can scale up to match the size of datasets now available. Here, we illustrate the scale at which a Learning Management System (LMS) accumulates data, and discuss the barriers to using this data for in-depth analyses. We illustrate the exponential growth of our LMS data to represent a single example dataset, and highlight the broader need for taking a pro-active approach to dimensional modelling in learning analytics, anticipating that common learning analytics questions will be computationally expensive, and that the most useful data structures for learning analytics will not necessarily follow those of the source dataset.

CCS CONCEPTS

• Social and professional topics ~ Management of computing and information systems • Information systems ~ Information storage technologies

KEYWORDS

Learning analytics, Learning Management Systems, data retention, big data, barriers to adoption, dimensional modelling

ACM Reference format:

S. Taylor and P. Munguia. 2018. Towards a data archiving solution for learning analytics. In *Proceedings of 8th International Learning Analytics and Knowledge conference, Sydney, Australia, March 2018* (LAK '18). DOI: 10.1145/3170358.3170415

1 INTRODUCTION

With the growth of digital data tools in the learning and teaching space, datasets of an unprecedented size are becoming available for use and analysis [1-4]. The potential of these large datasets to facilitate insights into learning and teaching processes has been recognised and anticipated now for over five years [1, 5, 6]. However, in practice, the use of educational data for analysis has lagged behind the growth in the size of datasets [5, 7, 8]. Taking a user-centred approach to learning analytics, we regard this in-practice gap between users and data as an important area for reflection.

The transition of "big" data to "meaningful" data has not been a simple process. Big datasets do not necessarily generate meaningful insights, and the assumption that larger datasets will necessarily be of higher quality, is incorrect. The 2015 Conference on Learning Analytics and Knowledge explored this topic in detail, with the theme of "Scaling Up: Big Data to Big Impact" [9]. The overlapping barriers to effective use of large educational datasets have been explored [5, 10]; these barriers include significant processing time and resource overheads to accessing and formatting large datasets.

Learning Management Systems (LMS), including software platforms such as Blackboard and Canvas produce large datasets. Learning Management Systems facilitate online delivery and management of education content. A sideline (and selling point) of these systems is that they enable tracking online activity from students and educators, and can deliver data derived from this activity [11].

Consistent with Kavanovic and colleagues [12] we contend that pre-processing is a necessary but underreported step in many learning analytics projects. Further, we contend that processing LMS data for use in learning analytics is in effect a dimensional modelling process, comparable to other subject-oriented data warehousing projects [13].

As an example, we demonstrate the exponential growth of timestamped online activity data in a newly implemented LMS at RMIT University. We wish to highlight that the dimensional models of datasets (big or small) can and should be optimised to suit the purposes to which they are applied, and that this is particularly important when datasets grow at a pace for which conventional data manipulation quickly becomes impracticable.

2 SCALING UP LEARNING ANALYTICS

2.1 Problem setting

Digital tools have significantly impacted the size and complexity of datasets available within educational settings. In particular, digital tools have greatly increased the amount of data retained: activities such as accessing files, viewing media, and interacting with discussions might otherwise have remained ephemeral, but are now accruing in datasets populated by digital tools.

LAK '18, March 7–9, 2018, Sydney, NSW, Australia
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-6400-3/18/03...$15.00
https://doi.org/10.1145/3170358.3170415

Big data sets including digital data tools in education began to accelerate in earnest after 2011, part of the broader suite of "big data" which gained the attention of both researchers and mainstream news at that time [14-17]. Reseachers were quick to see the analytical potential of abundant data sources in education [1, 6, 16] as well as the challenges [2, 3, 5, 18, 19].

In the years since "big data" entered discourse, researchers have noted that, in practice, analytical use of data in education has lagged behind the growth in the size of the datasets [5, 10, 20, 21]. Multiple overlapping social, operational, and technical constraints have been identified as barriers to use of large learning and teaching datasets for productive analysis [7, 10, 18, 21]. For example: a lack of clear ethics and privacy frameworks [22]; isolation of datasets into departments [5]; one-size-fits-all vendor solutions that cannot be modified by institutions [18]; and lack of partnership between IT departments and researchers, leading to access being driven more by technology than by policy [8]. Indeed, one of the major challenges is the access to data in the organisational complexity characteristic of universities [3, 23].

Bichsel [5] reported on a survey of educational institutions in the US, finding that while the majority of institutions regarded analytics as a priority, and were also collecting data, after four years the majority had also not progressed beyond a preliminary stage of data use. Similarly, in Australia learning analytics projects were reported as immature and of small scale [21], with many institutions associating the identification of students at risk with the entire field of learning analytics.

A dataset that is slow to access, complex to query, and difficult to integrate with other datasets, is not the only hurdle to analytical insights, but it is recognised as an exacerbating factor when combined with wider organisational complexities. Daniel [2], in a review of big data and analytics in higher education, identified two groups of analysts: those who know how to extract and identify available data, and those who understand which datasets are required by understanding analytical methods, with a divide still present between the two groups. This divide is a particularly important issue to consider as datasets grow larger and, concurrently, the value of integrating multiple datasets is increasingly recognised [9]. Without intervention and resourcing, large datasets rapidly become silos [5], in part because of the need for dataset specialists and analytical specialists to collaborate.

At the University of Sydney, Liu and colleagues [7] identified a disconnect between the offerings of algorithms and technology and the practical needs of instructors. This disconnect contributed to a destructive feedback loop, with inadequate tools (in spite of large datasets) contributing to lower interest in learning analytics, and thence to lower investment. They approached these issues by working on a bottom-up data delivery tool, delivering a relatively small but highly customised set of bespoke data to instructors. A key component of the effectiveness of the tool was a "high level of control over data collection and processing." [7]

2.2 How big is big data?

Big data typically refers to datasets with sizes in the order of Terabytes or Petabytes [14, 15]. In 2017, multiple companies process over 100 PB of data per day, hundreds of millions times the storage capacity of a personal computer in the mid-2000s. However, a more comprehensive definition of big data refers to the "three V's" of *volume*, *velocity*, and *variety*. One or all of these factors can make datasets difficult to manage even if they also, correspondingly, contain many potential insights.

Table 1 lists approximate dataset volume and velocity for several high-profile data-driven companies. It is important to note that a) no learning analytics dataset, including LMS platforms, approaches these volumes of data yet; b) nonetheless, all these companies make active use of data, and c) not coincidentally, none of these companies use data only in its raw form, and their processes for modelling and managing data are core pieces of intellectual property. With small teams and/or modest resources, researchers in learning and teaching venture into the use of datasets that are not the biggest but which can quickly exhibit high volumes and high complexity. As such, these projects can benefit from approaches to data already used in other fields of data management.

Table 1: Estimates of volume and velocity for contemporary "big data" companies [24]

Example dataset	Volume (data stored)	Velocity (data processed per day)
Google	15,000 PB	100 PB
Facebook	300 PB	600 TB
Twitter	(unknown)	100 TB
eBay	90 PB	100 PB
Spotify	10 PB	2.2 TB

2.3 A primer on data warehouses and dimensional models

Data availability is one of the main obstacles in learning analytics data, followed by not-quite-right data formats and delays in processing [16, 25, 26, 27].

These are near-textbook examples of reasons to consider creating a data warehouse [13]. A data warehouse is a specific type of database, where data are structured for the queries and analyses most commonly seen in a given organisation [28]. Data warehouses incorporate the ongoing work of extracting, transforming, and loading (ETL) data from multiple sources into a single landing place, where data is structured with the primary purpose of answering the most common questions of its most common users.

Data warehouses present a midway point between raw transaction data and perfect database normalisation, i.e. between a data structure that is both overly large and unuseable, and a data structure with the most efficient possible storage, which in practice is likely to be challenging and irritatingly slow to actually query (and so, in all probability, is also unlikely to be used). Instead of being structured into the smallest possible units, data warehouses are subject-oriented [29], built around customised *dimensional models* of *facts* and *dimensions*.

In dimensional modelling, facts describe a measurable quantitative item, while dimensions facilitate different ways to partition and identify facts, and can function as a vantage point from which to view facts. A fact table always includes a measurement field: e.g. time, cost, count. Dimensions are stored as rows in dimension tables; and as dimension key fields in fact tables.

Most importantly: the dimensional model most suited to a given purpose will not necessarily match that for another purpose, even if the source data is the same. The exact dimensional models used by large data-driven companies, such as those in Table 1, are not in the public domain. But it is safe to say that they do not store data in its raw form, nor in the same format as each other.

When data users find that they are continually building complex queries with the same components, with irritating time overheads at great computational expense, this can signal that dimensional modelling may be helpful. Dimensional data modelling therefore offers great benefits to learning analytics. Martin et al. [30, 31] showed the analytical benefits of working with dimensional data, with their research incorporating a data warehouse specifically for behavioral patterns of K-12 students. Similar benefits can be seen in the user-centred data structure of Liu, et al [7].

3 EXAMPLE DATASET

3.1 Canvas log data

Canvas is a cloud-based LMS released by Instructure, a Utah-based company and relative newcomer to the LMS marketplace.[1] Canvas includes a cloud-based system through which institutions can access daily "dumps" of their LMS data[2]. This is provided as a set of *fact tables* and *dimension tables*, as per the descriptions above.

Our current challenge is to build a useful landing place for this and other datasets, specifically for learning analytics projects. We wish to maintain the integrity of the source data schemas from which the data derives, but also to archive data in a place where it is accessible, useable, and able to be integrated with other datasets in a learning analytics environment. In considering the development of such a landing place, we are currently assessing the growth of Canvas data, and its most likely uses.

Our study timeframe covers the preliminary phase of introducing Canvas for a small fraction of the courses available at RMIT University. An instance of Canvas was initiated in August 2016. In Semester 2 2017, 50 courses were made active on Canvas with student enrolments, along with an additional set of several thousand "sandbox" courses for staff to become acquainted with the LMS user interface. Thus the study timeframe incorporates Canvas data for a very small subset of courses and student activity at our university.

[1] http://edutechnica.com/2017/09/17/5th-annual-lms-data-update/
[2] Canvas data structure documentation:
https://community.canvaslms.com/docs/DOC-10754
https://portal.inshosteddata.com/docs

In addition to the 82 fact and dimension tables, a *requests* table is made available for download. This table contains trace log data: timestamps and attributes of web requests sent to the Canvas LMS. The *requests* table is, strictly speaking, neither a dimension table nor a fact table. Trace log tables such as the Canvas *requests* table (or its Blackboard equivalent, the *activity accumulator* table) sit outside the dimensional model of the LMS, and are closer to raw usage data. They are provided by vendors to give institutions the opportunity to construct queries not already catered for by the LMS dimensional model. As such, they are popular with learning analytics projects (as described below, section 3.3). The *requests* table is the focus of our example growth rate charts below.

3.2 Volume and velocity

We have the opportunity to assess the volume and velocity of Canvas LMS data (Figure 1). The exponential growth of row counts for the *requests* log table with only 50 courses active with student enrolments is daunting; the record count in the *requests* table grew from 0 to 68.7 million records within one year. The record count doubled in the last month, from 28 million to 68.7 million records.

Although many tables grew considerably, the final record count in the *requests* table exceeded those from all fact and dimension tables combined: the 82 fact and dimension tables totalled 9.6 million records at the end of the year (15% of the total records in the *requests* table).

Similarly, the daily growth (velocity) of new records added to the *requests* table shows exponential growth (Figure 2). In the last month of the study timeframe an average of 974,000 new *request* records were added each day, with a maximum of 1.27 million new records in one day.

Without indexes, the *requests* table at the end of one year required 26 GB of disk storage in a PostgreSQL database, corresponding to 49 GB of disk storage with indexes[3]. This is well below the volumes of the big data companies listed in Table 1, but presents challenges to constructing useable queries. At this point, a simple grouping of user and course takes six minutes to execute; a non-trivial query would take considerably longer to build and to execute, and would be unuseable for any practicable user-focused reporting tool. In addition, the requirements to archiving this data in its present form would be unsustainable in the future given the velocity of incoming data.

3.3 What can these data be used for?

A selection of literature from Learning Analytics and Knowledge conferences, the *Journal of Learning Analytics*, and related sources provide examples of the types of metrics that are constructed from Learning Management Systems [12, 32-36]. *User* and *time* are consistent features. Researchers spend considerable time reconstructing log datasets like those in the

[3] Indexes and storage sizes vary by database configuration. Nonetheless, a database solution of some sort (whether cloud-based or on a local server) is necessary, given that opening this dataset, or querying it, is otherwise impossible. PostreSQL is also the native storage format of Canvas data.

Canvas *requests* table, to build time-related facts (duration, interval, sequence), and user-centred dimensions (e.g., user-and-task, user-and-instructor).

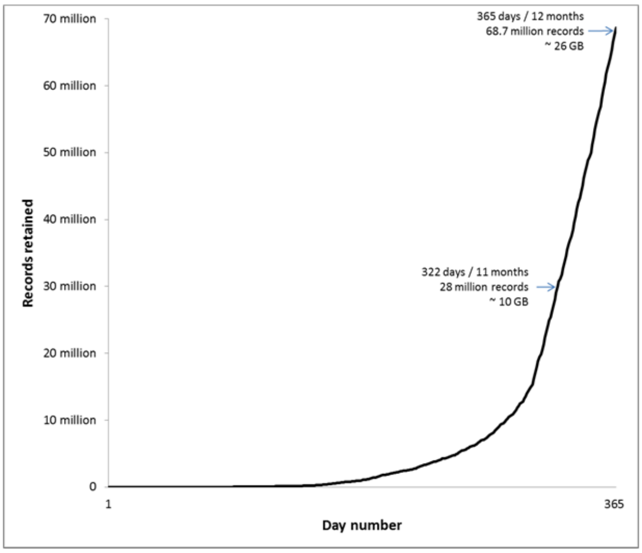

Figure 1: Number of records in the *requests* table, first year of implementation with only 50 courses live.

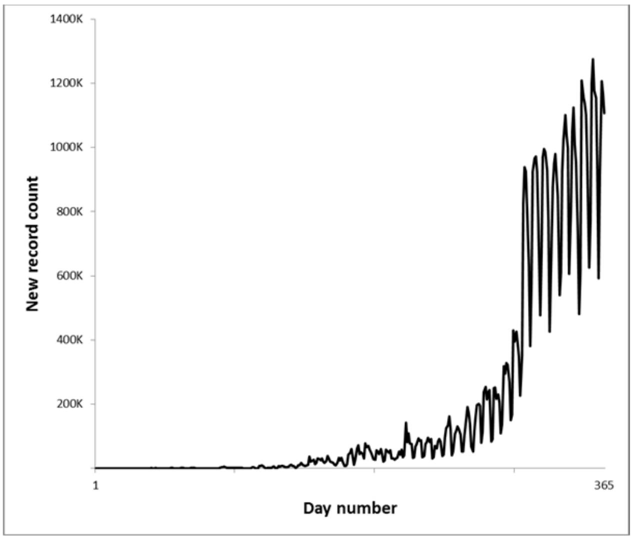

Figure 2: Daily growth of records by day for the *requests* table. Each dip is a weekend.

It should not be surprising, nor is it inherently problematic, that the datasets derived from Learning Management Systems are a) optimised for online web content management and not for the time-related facts and user-centred dimensions so often of interest within learning analytics, and b) very large. Repurposing this data for learning analytics is possible, but a non-trivial task.

Kovanovic and colleagues [12] reviewed the use of time-on-task analyses with LMS data. They noted that time-on-task measures are popular in spite of significant technical difficulties and inconsistencies. The authors note that in using LMS trace data, researchers are effectively embarking on Web Usage Mining (WUM), and that in any WUM task data pre-processing is expected to take 60% to 80% of total analysis time. Yet, they notes that processing overheads for time-on-task measures are rarely discussed within learning analytics [12].

By placing time as a primary dimension of interest, many analytical approaches to LMS data come to rely on tables such as *requests:* the trace log data is closer to raw form than the dimensional structure used by the LMS itself, and is often the only container for granular time data. Such tables are invariably the largest tables available in terms of row count, storage requirements, and query time, and come, essentially, with a "use at your own peril" stamp from vendors. Repurposing these datasets should involve a consideration of the facts and dimensions actually of interest. Analysing them in their raw form will soon become challenging, if not impossible.

3.4 Example dimensional modelling

In brief, the dimensional modelling process entails: a) identifying a process of interest; b) choosing a level of detail (granularity) for which an immediate reporting use exists; c) identifying the dimensions involved, and; d) identifying a measurement (fact). Large, granular fact tables offer flexibility at the expense of size and speed. We can experiment with different dimensionality combinations to understand resulting data sizes.

For example, in addressing questions such as *"when are students most active?"* or *"which courses are students most active in?"* one potential fact table from our example Canvas dataset can have two dimensions (*student* and *course*), with 17,000 records and 872 KB size. This table is small and useful for simple summaries, but with limited flexibility for exploratory analysis. An alternately structured fact table can contain three dimensions (*student*, *course* and *day*), with 231,000 records and 15MB size. This fact table offers greater flexibility but at a cost of slow query speed and large size. Both fact tables, however, are substantially smaller than the original *requests* table.

3.5 Limitations

The limitations of dimensional modelling are manifested in the literature presented in Section 2.1 [9, 10, 21]. There are two particular considerations to be wary of when embarking upon a dimensional modelling approach to learning analytics datasets. Firstly, just having access to data does not guarantee insights into data: any number of well-structured tables will not, in themselves, translate into greater use or insights. Secondly, the potentially daunting tasks and learning curves associated with these processes can create data silos, rigid processes, and lack of innovation: the very conditions that they are intended to mitigate. However, we contend that avoiding these concepts entirely is likely to have a similar effect. Hence, we simply argue for an awareness of dimensional modelling concepts within learning analytics.

4 CONCLUSIONS

In this paper we have provided an overview of the gap between size of datasets, and use of datasets, within learning analytics. We have introduced basic concepts of dimensional modelling, and drawn attention to the utility of these concepts within learning analytics. Our example dataset (the Canvas *requests* table) highlights the rate at which digital datasets in educational settings can grow to the point where common queries present significant challenges. We suggest that exponential data growth from big data sources such as Learning Management Systems can and should be pro-actively managed specifically for learning analytics.

In part, this means acknowledging that the facts and dimensions of most interest to learning analytics will not always match those from original data sources. But more importantly, data archiving solutions for learning analytics need to be dimensionally accessible for modelling purposes, rather than left to individual researchers to struggle with in isolation. The learning analytics community can make inroads to this by continuing to report honestly on data adoption barriers, and by utilising strategies already in use within wider data mangement fields. The Research Data Warehouse [31, 37] and the bespoke data delivery project at the University of Sydney [7] provide some working examples.

Given that learning and teaching data are already, and inevitably, placed within organisations characterised by great complexity, and that the datasets themselves are growing exponentially, failure to anticipate the volume, velocity and variety of incoming data, and the utility of dimensional modelling, can exacerbate one of the multiple barriers to productive analyses of available data. Pro-active strategies to provide a suitable landing place for learning analytics data can help to offset the downsides of big data, and are consistent with a user-centred approach to learning analytics.

REFERENCES

[1] M. Bienkowski, M. Feng, and B. Means. 2012. *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.* Washington, DC: US Department of Education, Office of Educational Technology.
[2] B. Daniel. 2015. Big data and analytics in higher education: Opportunities and challenges. *British journal of educational technology*, 46, 5 (2015), 904-920.
[3] L. P. Macfadyen, S. Dawson, A. Pardo, and D. Gasevic. 2014. Embracing big data in complex educational systems: The learning analytics imperative and the policy challenge. *Research & Practice in Assessment*, 9 (2014), 17-28.
[4] R. J. Vatsala, and R. Sathyaraj. 2017. A Review of Big Data Analytics in Sector of Higher Education. *International Journal of Engineering Research and Applications*, 7, 6 (2017), 25-32.
[5] J. Bichsel. 2012. *Analytics in higher education: Benefits, barriers, progress, and recommendations.* Louisville, CO: EDUCAUSE Center for Applied Research.
[6] J. P. Campbell and D. G. Oblinger. 2007. Academic analytics: A new tool for a new era. *EDUCAUSE review*, 42, 4 (2007), 40-57.
[7] D. Y. Liu, C. E. Taylor, A. J. Bridgeman, K. Bartimote-Aufflick and A. Pardo. 2016. Empowering Instructors Through Customizable Collection and Analyses of Actionable Information. *Proceedings of the First Learning Analytics for Curriculum and Program Quality Improvement Workshop, Hosted by the Learning Analytics and Knowledge (LAK).* ACM, Edinburgh, UK, April 25-29 (2016), 3-9.
[8] S. Lonn, S. Aguilar and S. D. Teasley. 2013. Issues, challenges, and lessons learned when scaling up a learning analytics intervention. *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK).* ACM, Leuven, Belgium, April 2013.
[9] A. Merceron, P. Blikstein and G. Siemens. 2016. Learning analytics: from big data to meaningful data. *Journal of Learning Analytics*, 2, 3 (2016), 4-8.
[10] R. Ferguson, D. Clow, L. Macfadyen, A. Essa, S. Dawson and S. Alexander. 2014. Setting learning analytics in context: overcoming the barriers to large-scale adoption. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge (LAK).* ACM, Indianapolis, IN, USA, March 24-28 (2014), 251-253.
[11] W. Watson and S. L. Watson. 2007. An Argument for Clarity: What are Learning Management Systems, What are They Not, and What Should They Become. *TechTrends*, 51, 2 (2007), 28-34.
[12] V. Kovanovic, D. Gašević, S. Dawson, S. Joksimovic and R. Baker. 2016. Does time-on-task estimation matter? Implications on validity of learning analytics findings. *Journal of Learning Analytics*, 2, 3 (2016), 81-110.
[13] R. Kimball and M. Ross. 2011. *The data warehouse toolkit: the complete guide to dimensional modeling.* Second edition. New York, NY: John Wiley & Sons.
[14] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh and A. H. Byers. 2011. *Big data: The next frontier for innovation, competition, and productivity.* McKinsey Global Institute, May 2011.
[15] C. Howson. 2013. *Successful business intelligence: Unlock the value of BI & big data.* Second Edition. McGraw-Hill Education Group.
[16] G. Siemens and P. Long. 2011. Penetrating the fog: Analytics in learning and education. *EDUCAUSE review*, 46, 5 (2011), 30-32, 34, 36, 38, 40.
[17] S. Lohr. 2012. The Age of Big Data. *New York Times*, February 11 2012. www.nytimes.com/2012/02/12/sunday-review/big-datas-impact-in-the-world.html
[18] D. Y-T. Liu, T. Rogers and A. Pardo. 2015. Learning analytics – are we at risk of missing the point? *Proceedings of Australasian Society for Computers in Learning in Tertiary Education (ascilite).* Perth, Australia, November 20 - December 2 (2015), 684-687.
[19] R. Ferguson. 2012. Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4, 5-6 (2012), 304-317.
[20] W. Greller and H. Drachsler. 2012. Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*, 15, 3 (2012), 42.
[21] Colvin, C., Rogers, T., Wade, A., Dawson, S., Gašević, D., Buckingham Shum, S. and J. Fisher. 2015. *Student retention and learning analytics: A snapshot of Australian practices and a framework for advancement.* Sydney, Australia: Australian Office for Learning and Teaching.
[22] G. Siemens, S. Dawson and G. Lynch. Improving the quality and productivity of the higher education sector. *Policy and Strategy for Systems-Level Deployment of Learning Analytics.* Canberra, Australia: Society for Learning Analytics Research for the Australian Office for Learning and Teaching.
[23] M. D. Cohen and J. G. March. 1974. *Leadership and ambiguity: The American college president.* New York: McGraw-Hill.
[24] M. Huss and J. Westerberg. 2014. Data size estimates. *Follow the Data.* June 24 2014. https://followthedata.wordpress.com/2014/06/24/data-size-estimates/
[25] C. Beer, K. Clark and D. Jones. 2010. Indicators of engagement. *Proceedings of Australasian Society for Computers in Learning in Tertiary Education (ascilite).* Sydney, Australia, December 5-8 (2010), 75-86.
[26] S. Dawson, L. Heathcote and G. Poole. 2010. Harnessing ICT potential: The adoption and analysis of ICT systems for enhancing the student learning experience. *International Journal of Educational Management*, 24, 2 (2010), 116-128.
[27] J. Ma, X. Han, J. Yang and J. Cheng. 2015. Examining the necessary condition for engagement in an online learning environment based on learning analytics approach: The role of the instructor. *The Internet and Higher Education*, 24 (2015), 26-34.
[28] R. Kimball. 1996. *The data warehouse toolkit: practical techniques for building dimensional data warehouse.* New York: John Willey & Sons.
[29] W. Harrison and H. Ossher. 1993. Subject-oriented programming: a critique of pure objects. *Proceedings of the eighth annual conference on Object-oriented programming systems, languages, and applications.* Washington, DC, USA, September 26-October 1 (1993), 411-428.
[30] D. C. Nacu, C. K. Martin, N. Pinkard and T. Gray. 2016. Analyzing educators' online interactions: a framework of online learning support roles. *Learning, Media and Technology*, 41, 2 (2016), 283-305.
[31] D. Nacu, C. K. Martin, M. Schutzenhofer and N. Pinkard. 2016. Beyond traditional metrics: Using automated log coding to understand 21st century learning online. *Proceedings of the Sixth International Conference on Learning Analytics And Knowledge (LAK).* ACM, Edinburgh, UK, April 25-29 (2016), 197-200.
[32] L. P. Macfadyen and S. Dawson. 2010. Mining LMS data to develop an "early warning system" for educators: A proof of concept. *Computers & education*, 54, 2 (2010), 588-599.
[33] T. M. Khan, F. Clear and S. S. Sajadi. 2012. The relationship between educational performance and online access routines: analysis of students' access to an online discussion forum. *Proceedings of the Second International Conference on Learning Analytics and Knowledge (LAK).* ACM, Vancouver, Canada, April 29-May 2 (2012), 226-229.
[34] M. Andergassen, G. Neumann and F. Mödritscher. 2013. The four seasons: Identification of seasonal effects in LMS usage data. Villard-de-Lans, France, January 28-February 1 (2013).

[35] J. Fritz. 2016. LMS Course Design As Learning Analytics Variable. *Proceedings of the First Learning Analytics for Curriculum and Program Quality Improvement Workshop, Hosted by the Learning Analytics and Knowledge (LAK).* ACM, Edinburgh, UK, April 25-29 (2016), 15-19.
[36] T. Hecking, S. Ziebarth and H. U. Hoppe. 2014. Analysis of dynamic resource access patterns in online courses. *Journal of Learning Analytics*, 1, 3 (2014), 34-60.
[37] C. K. Martin, D. Nacu and N. Pinkard. 2016. An Approach to Using Log Data to Understand and Support 21st Century Learning Activity in K-12 Blended Learning Environments. *Journal of Learning Analytics*, 3, 2 (2016), 37-87.