Observing the Use of e-Textbooks in the Classroom: Towards "Offline" Learning Analytics

Maka Eradze, Terje Väljataga, and Mart Laanpere(✉)

Institute of Informatics, Tallinn University, Narva mnt 25, 10120 Tallinn, Estonia
{Maka.Eradze, Terjev, Mart.Laanpere}@tlu.ee

Abstract. Learning analytics is an emerging approach that is equally popular among researchers and educators-practitioners. Although the methods and tools for LA have been developing fast, there still exist several unsolved problems: LA is too much data driven, weakly connected to theory and is able to analyse only the activities documented in an online setting - in LMS. We propose a solution for the LA unit of analysis drawing upon the research of existing practices and tools used for offline contexts: the data is coming from the physical learning interactions based on the observations in the classroom setting and captured with classroom observation application. We argue that if the unit of analysis has a particular logic and structure, it can unleash the possibilities for "offline" analytics that can be later integrated with online LA.

Keywords: LEARMIX · Learning analytics · eTextbooks · Unit of analysis · TinCan API · xAPI

1 Introduction

Textbooks have been playing an important part in teaching and learning in the formal education context for more than one century. As the textbook publishers, editors and authors are the most careful readers and implementers of curricula and subject-related news, the textbooks have gained large impact in educational development. On the other hand, widespread use of printed textbooks is hindering the advancement of modern learning analytics, as the learning activities that take place outside of digital realm leave no digital trace behind. Emerging e-textbooks might change a lot in that sense, but it depends on the approach taken in e-textbook development. Currently, the majority of e-textbooks are released either as e-books (in epub, mobi, pdf formats), apps or content packages integrated into online course (e.g. SCORM or CommonCartridge). In most of the cases, only the latter format allows acquiring rich data about learning interactions in a standardized format. LEARNMIX project in Tallinn University aims at exploring alternative forms for e-textbooks of the future that should support innovative pedagogical scenarios and advanced learning analytics. But even if most of the textbooks and other learning resources will be turned into digital format, there will remain many learning activities that will take place in the physical classroom setting without leaving any digital trace behind for learning analytics. The research problem addressed by this

Y. Cao et al. (Eds.): ICWL 2014 Workshops, LNCS 8699, pp. 254–263, 2014.
DOI: 10.1007/978-3-319-13296-9_28

exploratory study was to find out the existing approaches and tools for collecting learning analytics data in the offline settings.

2 Unit of Analysis in Education and Computer Supported Collaborative Learning

When addressing research questions, it is important to have a consistent and theory driven unit of analysis. The discussion on the different *units of analysis,* approaches and developments throughout centuries and the philosophical stances they take on is important when it comes to learning analytics and its *unit of analysis.*

Educational research concentrates on different *units of analysis*; Stahl [1] discusses the issue of *unit of analysis* in cognition that had different foci in different times: *concepts* (Plato), *mental and material objects* (Descartes) (and relationship between them), *observable physical objects* (empiricism), *mind's structuring categorization efforts* (Kant). All of the approaches dealt with the inner functions of an individual mind. Hegel entered the discussion with a larger unit of analysis – which was historically, socially and culturally determined.

Hegel's philosophy shaped three mainstream schools of thought – Marx (critical social theory), Heidegger (existential phenomenology) and Wittgenstein (linguistic analysis). To Stahl, these three main directions influence how the CSCL units of analysis are formed: For Heidegger the unit of analysis was the *man with unified experience of being-in-the-world.* Wingenstein entered the discussion with the *unit of analysis* from *mental meanings* to *interpersonal communications in the context of getting something done together.*

In some cases CSCL research takes socio-cognitive or socio-cultural approaches. But in both cases the *unit of analysis* is mostly *an individual mind*. Engestrom is the one taking the unit of analysis to the whole *activity system*. But to Stahl's understanding Engestrom's theory is not interested in group knowledge building but rather with organizational management of the group. Influenced by Marx, theory tries to see societal issues in the making. Even in distributed cognition, which deals with group-cognitive phenomena, mostly socio-technical systems and highly developed artifacts are analyzed [1].

3 Learning Analytics: The Concept, the State of the Art

One of the leading definitions of learning analytics suggests that it is *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*. This definition had been set out at the 1st International Conference on Learning Analytics and Knowledge [2]. The field is still emerging, rapidly developing and experiencing *a gradual shift away from technology towards an educational focus*, while the three main drivers for learning analytics have been defined as technological, pedagogical and political/economic [3].

These drivers are conceptualized by Ferguson as challenges [3]:

- Big Data - a challenge for its volume, difficulty to handle the interaction data and most importantly extracting value from the big data-sets.
- Online Learning that poses an educational challenge - how to optimize opportunities for online learning.
- Political Concerns - how to improve learning opportunities and results at different levels?

According to Ferguson the drivers draw attention to the three groups of interest - governments, educational institutions and teachers/learners. The development of learning analytics shifts the balance between the three drivers and three groups.

Greller and Draschler give a general framework of learning analytics [4] and offer considering six critical dimensions within the research lens. Each of the dimensions can have several values and it can be extended upon a need. Represented dimensions are: stakeholders, objectives, data, instruments, external constraints and internal constraints.

Greller and Draschler also give a model of information flow between the stakeholders and it is based on a common hierarchical model of the formal education. A pyramid view (with the learner as a cornerstone) is illustrating how data analysis from lower layer can inform the above layer. According to Buckingam Shum the convergence of the Macro, Meso and micro levels is the key to the successful learning analytics [5].

3.1 Units of Analysis in LA

When we are to consider what has to be analyzed and what information do we need to infer using LA, firstly, the level of learning analytics must be defined. The interest groups may overlap but different granularities are needed for different groups: *The choice of target audience affects how researchers conceptualize problems, capture data, report findings, predict what will happen, act on their findings and refine their models* [3]. Within the context of our research interest the micro-level, teacher/learner learning analytics should be directed to the activity, an event consisting of interaction between **a subject** and **an object** that are bound with **a verb.** There is a need for theory driven, event oriented unit of analysis [6, 7].

Suthers et al. with the uptake framework proposed that the event is the core for analyzing data and understanding which interactions lead to learning [9]. The *Uptake Framework* [7, 9] assumes that interaction is fundamentally relational, so the most important unit of analysis is not isolated acts, but mostly relationships between acts.

Conceptualizing the Uptake Framework hierarchies and the possibilities of learning analytics, it has also been suggested to view Learning Flow as a main unit of analysis for the learning interaction analysis [7].

3.2 Limitations of LA and Potential of xAPI

Most of the tools for gathering the learning analytics data are directed to the closed LMS systems, while the most of the learning happens outside the LMS - in distributed setting or offline part of the learning which is most of learning. Currently, LA covers only the part of the learning that happens within the LMSs. In most of the cases, LMSs data is harvested and analyzed. The problem is that it is not enough. Siemens believes, that LMSs are adopted as learning analytics tools and reflect the learner's interactions within a system. The capabilities of tracking and visualisation of interaction data has also been limited [3, 8].

The similar problem persists with the physical world i.e. offline "data" - library uses, learning support, in case of blended learning - the part of the learning that happens outside of LMS, online or offline. Long and Siemens suggest mobile devices as prospects of "*bridging the divide between the physical and digital worlds*" [8].

One way of dealing with the limitations of leveraging the data from the settings outside LMS is to explore potential of Experience API [10]. The Experience API is a service that allows for statements of experience (typically learning experiences, but could be any experience) to be delivered to and stored securely in a Learning Record Store. Learning activity is a unit of instruction, experience or performance that has to be tracked. A Statement consists of <Actor (learner)> <verb> <object>, with <result>, in <context> to track an aspect of a learning experience. Several statements can be used to track the whole experience. The statements are recorded in the LRS - Learning Record Store [10].

Another problem with learning analytics within the limits of the current development is a weak connection to theory. This limitation of data monitoring and harvesting could be overcome by having a particular theory in mind before recording the data [7].

Our paper targets the "offline" analytics dilemma and explores the potential of xAPI and Uptake Framework working together towards a new type of *unit of analysis* in the context of learning analytics.

3.3 Ethical Considerations

It should not be argued that the privacy of the data subjects must be protected. There are several factors influencing the process of protection that can work against individual freedoms (if privacy is abused) or restrict using the full potential of LA. We believe these two factors shall be balanced. According to Hilderbrandt [11], the core of privacy must be found itself in the idea of identity and this is not only because of the advancement of high-tech identification technologies but also because the process of identity building can harm the privacy of individuals.

Slade et al. [12] believe that students shall be involved in the data harvesting and analysis. According to Kruse et al. [13] there should be a "student-centric", as opposed to an "intervention-centric", approach to learning analytics. This suggests the student should be seen as stakeholders of their own data. And also as co-interpreters of own data - and perhaps even as participants in the identification and gathering of that data. Greller et al. [4] list the ethical side of the use of personal data in the external limitations of learning analytics.

Based on the literature overview, currently we may refer to some of the solutions for data privacy protection: 1. Involving students [data subjects] in the process, make it transparent and make it a student analytics. 2. Anonymization/deidentification of data. 3. Consent forms.

4 eTextbook Analytics

4.1 Studies on eTextbooks Use

According to Baek et al. [14] *in order to effectively support students' learning, it is important to comprehend students' experiences using eTextbooks*. There are several possibilities to understand the patterns of use for future inferences – 1. For the deployment of appropriate pedagogic strategies 2. For student self-reporting 3. For decision making processes – in terms of the design and etc.

Research on the use of eTextbooks mainly focus on the issues of satisfaction of use, preference of use over traditional textbooks and other factors [15–18]; The study conducted by Baek et al. [14] in the various campuses of US focuses on the understanding of students' eTextbook use experiences. This study used surveys to assess students' perceptions of the eTextbook in terms of satisfaction and ease of use. Cutshall et al. [17] also assessed perceptions on the use of etextbooks and web-based homeworks. When assessing the use of eTextbooks logs were only used to understand student reading behaviors (number of page prints) and correlated to the satisfactions of use [18].

An example of analytics in eTextbooks is a research conducted by Nicholas et al. [19] using the data from digital footprints on (a) volume, duration and timing of use; (b) where use took place; (c) individual book titles used; (d) location of use; (e) type of page viewed; (f) institutional and subject diversity; (g) scatter of use; (h) nature of use; and (i) method of searching/navigating. The log data were analysed to describe how users interacted with the system. The authors, though, conclude that *logs only provide us with a very superficial idea of who the e-book users were (their institutional affiliation was known), so for a better picture we have to turn to the questionnaires.* Khurana et al. [20] deployed text analytics to assess the coverage, readability and comprehensibility of eTextbooks. They use different units of analysis: sections, bookmarks, topics, sub-topics.

Having a goal to build an open source online eTextbook for DSA courses integrating textbook quality text with algorithm visualization and interactive exercises, Fouh et al. [21] concentrate on the development of a OpenDSA interactive eTextbook where they also incorporate a kind learning analytics – mainly for the self-reporting for students, and also inferring meaning from student-content interactions for "studying the pedagogical effectiveness for various approaches and support for gathering data about usability of system components for future improvement. So the unit of analysis is mainly student-content interaction centered. The study on the use of the eTextbook was aimed at the student perceived satisfaction evaluation and a test whether the eTextbook helped reduce the grading burden.

Studies on eTextbook use are developed around the ideas of satisfaction of use or reading behaviors. *Units of analysis* are individual student perceptions and sometimes student-content interactions to gain insight on reading behaviors, not the analysis of the design or pedagogical rationale behind it. Very often, when the study aims at uncovering the learning design principles of an eTextbook, it does not refer to the possibilities of learning analytics as for instance, in case of the study of Choi et al. [22].

4.2 Offline Learning Analytics: Observing the Use of Textbooks in the Classroom Lesson Observation Apps: Critical View

Two different approaches can be used in the eTextbook use observations: taking advantage of online data coming from clicks, resource access etc. and "offline" analytics with its wide range of possible interactions, written in different statements and formats.

Classroom observation apps are very useful tools for recording classroom learning interactions on the use of textbooks in "offline' settings. For this particular observation study we overviewed and compared 6 classroom observation apps based on particular requirements. These applications are: LessonNote, iObserve, Observation 360, iAspire, GoObserve, SCOA. Applications were chosen according to their free access to at least demo versions.

The applications were compared considering several features: 1. Interface affordances 2. The ways of input 3. Pedagogic scenario/model 4. Output of the generated data 5. Possibilities of analytics and most important part of our research scope 6. Units of analysis. The features were chosen based on the importance to the scope of the research. The table describes the proportion of certain features used in those applications (Table 1).

Table 1. Application comparison

Feature	Value					
Interface	Tapping .6	Drag&drop .1	Sliders .3			
input	Handwritting .09	Typing .36	Photo .27	Audio .09	Video .09	Other .09
Scenario/ model	Based on a spec.model .5	Based on several models .33	Flexibility of switching models 0	Neutral .17		
Output	.pdf .25	Cvs .0	Word .08	Email .42	database/ cloud .25	
Analytics	No analytics .31	No datasets .15	visualisations 0.31	datasets/ cloud .3		
Unit of analysis	individual/ teacher .43	individual/ student .14	Event/activity .07	Group .07	Class .29	individual/ teacher .43

Based on the overview LessonNote app was chosen for it represents the closest possible app to what we have envisioned for the use in observations, namely for its event-driven *unit of analysis*.

5 Empirical Study

In the remaining part of the paper we will describe our effort to use a LessonNote application for supporting the collection of offline learning analytics while observing the use of textbooks in the classroom settings. We will continue with analysing and demonstrating benefits and drawbacks of LessonNote application for recording offline learning analytics. The study mainly focuses on the *unit of analysis* and its importance in the “offline” analytics based on the classroom observation application.

5.1 Method and Sample

In the context of Learnmix project we carried out an intervention study in K-12 education. Our aim was to intervene into current teaching and learning practices with the purpose to enable learners to become actively engaged constructors of their own experience and knowledge by creating, modifying and integrating various physical, and digital artefacts. For that we designed five different scenarios (flipped- classroom, project-based learning, game-based learning, inquiry-based learning, problem-based learning) for teachers to choose from and implement it in her/his lessons. In these scenarios the role and use of textbooks changed from textbooks as an object of knowledge construction to textbooks as a source of inspiration, etc. We have to mention here that we do not treat the aforementioned list of scenarios as a definite one, but rather as a starting set of potential scenarios for enabling students to become constructors of their own experience and knowledge in the midst of the digital transformation.

We observed 12 lessons in 6 different K-12 schools. These schools were chosen because of their more advanced IT infrastructure and teachers with open-minded learning and teaching practices. For documenting the flow of a lesson and emerging interactions we made use of LessonNote application. LessonNote application allowed timing, recording photos of student work and activities, which were inserted into the notes; and creating seating charts. As an additional tool we video recorded all the observed lessons. For the research described in this paper the videos didn’t play an essential role.

For understanding the use of (e-)textbooks in the aforementioned scenarios we created a framework for extracting the statements of students and teachers’ experience (learning flow) in a similar way to Experience API. Our framework consists of three main items:

1. Actors - a teacher or student(s) specifying whether the activity was done in groups, peers or individually.
2. Artifacts - artifacts were divided into three groups:

   – Display artifacts are physical objects in the classroom (for instance computer, projector, screen) whose function is to display conveyor artifacts. Display artifacts themselves are not representations of knowledge, but are seen as carriers for other artifacts.
   – Conveyor artifacts are various applications, which support the mediation or creation of knowledge representations (for instance iBooks, Prezi, Weblog, etc.). The affordances of conveyor artifacts very often define potential actions.
   – Content artifacts are representations of knowledge displayed in different formats (for instance text, video, image), which are created by professional textbook authors, by teachers, by students or others.

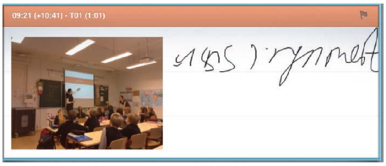
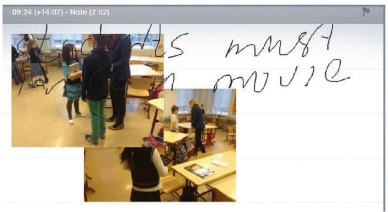
3. Actions - actions performed by a teacher or student(s) during the learning experience.

Such a framework allowed us to focus on specific actions and every accompanying (digital) artifact used or created before and during the learning experience. Furthermore, for our intervention study it was important for us to determine the role of students and teachers in learning experiences (whether a teacher or a student is a creator of an (digital) artifact, whether a student takes control and responsibility for what he/she is doing, etc.).

5.2 Results and Discussion

We implemented our analytical framework to our data set extracted from LessonNote application and video transcript. Despite of its many useful affordances, such as allowing recording activities according to timeline, shooting photos and adding them to a particular activity, LessonNote application also has some deficiencies. With the following 2 examples we demonstrate the deficiencies of LessonNote application as a tool for supporting the collection of offline learning analytics and translating its data into a form that supports Experience API statements and Uptake framework (Table 2).

Table 2. Results

Activities	**LessonNote activities**	**Video transcript: (Subject verb object)**
Teacher activity		Teacher organizes class **Teacher gives assignment** Teacher forms groups
		Group A moves out **Group A organizes tools** Group A starts a discussion

In the table we presented examples from the LessonNote app aligned with data coming from video transcripts. Video transcripts were produced by two researchers putting in the matrix compatible with xAPI statements. The examples brought here demonstrate how the LessonNote app captured activities and what can be extracted from videos. LessonNote captures one particular activity (shown in bold) and with video and later analysis it is possible to capture *preceding* and *proceeding* activities with the LessonNote captured activity encapsulated by the two (and more). But also this is to show that it is possible to structure the data in the form of Experience API compatible statements.

5.3 Conclusion and Future Work

The intervention study showed that it is possible to transcribe the interaction data in the form of statements, but recording “offline” interactions with LessonNote app did not offer satisfactory results for several reasons:

1. It proved to have interface problems – it is not possible to handwrite data as it is happening in real time.
2. It does not capture nested activities.
3. It does not allow quick documentation of activities.
4. It has no enough affordances, for instance it is not possible to define/form groups and assign numbers for later analysis.
5. Though it more or less focuses on event as a unit of analysis, it does not give full possibilities to automatize the process.
6. It does not show the dyadic interactions - who is interacting with whom.

Based on the overview of classroom applications and the empirical study we plan to develop a classroom observation application to be used on offline observations and learning analytics. This application will cover the gaps and offer “offline” analytical features that can potentially be aligned with online data. The application will be based on the overview of the similar applications and xAPI statement and event-driven *unit of analysis.*

References

1. Stahl, G.: Theories of cognition in collaborative learning. In: Hmelo-Silver, C., O’Donnell, A., Chan, C., Chinn, C. (eds.) International Handbook of Collaborative Learning. Taylor & Francis, New York (2012)
2. Learning Analytics & Knowledge, Banff, Alberta, February 27–March 1 2011. https://tekri.athabascau.ca/analytics/
3. Ferguson, R.: The state of learning analytics in 2012: a review and future challenges. Technical report KMI-12-01, Knowledge Media Institute, The Open University, UK (2012)
4. Greller, W., Drachsler, H.: Translating learning into numbers: a generic framework for learning analytics. Educ. Technol Soc. **15**(3), 42–57 (2012)
5. Buckingam, S.S.: Learning Analytics. UNESCO Policy Brief, Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief U.S. Department of Education Office of Educational Technology (2012)
6. Barab, S.A., Evans, M.A., Baek, E.O.: Activity theory as a lens for characterizing the participatory unit. In: Jonassen, D.H. (ed.) Handbook of research on educational communications and technology, pp. 199–214. Lawrence Erlbaum Associates, Mahwah (2004)

7. Eradze, M., Pata, K., Laanpere, M.: Analyzing learning flows in digital learning ecosystems. In: Huang, Y.-M., Li, F., Jin, Q. (Toim.) Knowledge Management and E-learning. LNCS, pp. 1–10. Springer, Heidelberg (in press)
8. Long, P., Siemens, G.: http://www.educause.edu/ero/article/penetrating-fog-analytics-learning-and-education (2011). Accessed 20 June 2014
9. Suthers, D. D., Rosen, D.: A unified framework for multi-level analysis of distributed learning. In: Proceedings of the First International Conference on Learning Analytics and Knowledge, pp. 64–74. ACM, Banff, New York (2010)
10. Experience API Specification. http://tincanapi.wikispaces.com/file/view/Experience+API+Release+v0.95.pdf. Accessed 20 June 2014
11. Hildebrandt, M.: Privacy and identity. In: Claes, E., Duff, A., Gutwirth, S. (eds.) Privacy and the Criminal Law, pp. 61–104. Intersentia, Antwerp/Oxford (2006)
12. Slade, S., Prinsloo, P.: Learning analytics: ethical issues and dilemmas. Am. Behav. Sci. **57** (10), 1509–1528 (2013)
13. Kruse, A., Pongsajapan, R.: Student-Centered Learning Analytics. https://cndls.georgetown.edu/m/documents/thoughtpaper-krusepongsajapan.pdf (2012). Accessed 20 June 2014
14. Baek, E., Monagham, J.: Journey to textbook affordability: an investigation of students' use of eTextbooks at multiple campuses. Int. Rev. Res Open Distance Learn. [S.l.] **14**(3), 1–26 (2013). ISSN 1492-3831. http://www.irrodl.org/index.php/irrodl/article/view/1237. Accessed 6 June 2014
15. Weisberg, S.: Student attitudes and behaviors towards digital textbooks. Publishing Res. Q. **27**, 188–196 (2011). Springer
16. Dennis, A.: e-Textbooks at Indiana University: A Summary of Two Years of Research. http://etexts.iu.edu/files/eText%20Pilot%20Data%202010-2011.pdf. Accessed 6 June 2014
17. Cutshall, R.C., Mollick J.S., M. Bland E.M.: Use of an e-Textbook and web-based homework for an undergraduate business course: students' perceptions. J. Acad. Bus. Educ. **10** (2009)
18. Falc, E.O.: Assessment of college students' attitudes towards using an online e-textbook. Interdisc. J. E-learn. Learn. Objects **9**, 1–12 (2013)
19. Nicholas, D.: E-textbook use, information seeking behaviour and its impact: case study business and management. J. Inf. Sci. **36**(2), 263–280 (2010)
20. Khurana S., Relan, M. Singh, V.K.A.: Text analytics-based approach to compute coverage, readability and comprehensibility of eBooks. In: 2013 Sixth International Conference on Contemporary Computing (IC3) (2013)
21. Fouh, E., Karavirta, V., Breakiron, D.A., Sally Hamouda, S., Hall, S., Naps, T.L., Shaffer, C. A.: Design and architecture of an interactive eTextbook – the OpenDSA system. Sci. Comput. Program. **88**, 22–40 (2014)
22. Choi, J.-I., Heo, H., Lim, K.Y., Jo, I.-H.: The development of an interactive digital textbook in middle school English. In: Kim, T.-h, Adeli, H., Slezak, D., Sandnes, F.E., Song, X., Chung, K.-i, Arnett, K.P. (eds.) FGIT 2011. LNCS, vol. 7105, pp. 397–405. Springer, Heidelberg (2011)