SOLAR
SOCIETY for LEARNING ANALYTICS RESEARCH

Creating Data for Learning Analytics Ecosystems

Learning is a complex process that involves rich interactions between people, politics, places, and increasingly, technology. Using clickstream data to provide deep insights into learning requires care and a system wide approach. We need learning analytics ecosystems.

- Kirsty Kitto, Connected Intelligence Centre, University of Technology Sydney
- John Whitmer, Federation for American Scientists
- Aaron E. Silvers, Elsevier Inc.
- Michael Webb, Jisc

September 2020

About SoLAR

The Society for Learning Analytics Research (SoLAR) is an interdisciplinary network of leading international researchers who are exploring the role and impact of analytics on teaching, learning, training and development.

SoLAR has been active in organizing the International Conference on Learning Analytics & Knowledge (LAK) and the Learning Analytics Summer Institute (LASI), launching multiple initiatives to support collaborative and open research around learning analytics, promoting the publication and dissemination of learning analytics research, and advising and consulting with state, provincial, and national governments. SoLAR's official publications in addition to LAK are the Journal of Learning Analytics and the Handbook of Learning Analytics.

Learn more and join the community at https://www.solaresearch.org

SoLAR Position Papers

SoLAR Position Papers are designed to catalyse debate around key opportunities that could advance the field of Learning Analytics, if ways to tackle the associated challenges can be developed.

The 2011 Position Paper envisioning Open Learning Analytics, published in the early days of the community, has served as a reference point since for imagining how the field might develop.

We warmly invite you to share your thoughts on this Position Paper via the dedicated thread in the Learning Analytics discussion forum.

Executive Summary [General Comments on Whole Paper]

Data and analytics infrastructure is starting to achieve a mainstream status in education, where many institutions believe that it can provide a foundation for extracting deep insights into numerous ways that we might start to improve the student journey. Accordingly many vendors in the Educational Technology (EdTech) landscape are now investing significant resources into the design and implementation of analytics products. This trend intersects with broader discussions around the value of data to society - where metaphors of data as "the new oil" are common, but increasingly disputed. However, data can only become valuable with careful extraction, inspection, and processing. Claims that data standards solve this problem have proven to be overly optimistic. As practitioners who have spent a number of years grappling with educational data, from both within institutions and industry perspectives, we use this position paper to draw attention to the issues that we have faced. In short: our EdTech systems are currently generating large amounts of low quality data that, while providing surface level insights, offers few avenues for improving student learning and the environments in which it occurs. We argue that this is due to a historical accident, where much of the data provided in EdTech systems was originally created for other purposes, such as debugging the systems themselves. Far more intentional planning must start to take place about how this data is created and stored before we can start to generate data that has potential for improving student learning. Collecting digital exhaust from teachers and students is both: unlikely to lead to better outcomes; and likely to result in inaccurate or misleading conclusions about our students. There is a risk that the current status quo will lead to decision makers and other stakeholders questioning the value of their investments in data and analytics, and starting to lose interest in the field as it fails to live up to their expectations. We examine the underlying reasons for these issues and provide recommendations to help institutions, vendors and practitioners working in the fields of LA/EDM, to ensure that they can deliver useful data for learning analytics ecosystems.

Contents

Introduction

Institutions, schools and the professional training sector all seek to improve learning outcomes for their students. Increasingly, organisations are starting to use data to infer insights that can help them achieve this aim. For many years, a number of student characteristics (e.g. course grades, learning competencies, feedback about teaching, and eventual career success) have been analyzed to derive insights about successful pedagogical approaches. However, these approaches were missing a critical component: the learning behaviors of students themselves. As educational technology (EdTech) becomes more integrated with the student learning experience, and modern data science increasingly starts to influence EdTech, this behavioral data is becoming a hot resource, but how useful is it really?

Aiming to develop insights from this new source of educational data, fields like Learning Analytics (LA) and Educational Data Mining (EDM) have grown dramatically in their popularity during the past few years. Accordingly, institutions are starting to include data requirements when purchasing products and services for their EdTech ecosystems. This has resulted in many EdTech vendors moving towards models where they aim to deliver data and analytics as a key feature of their applications, and even to position themselves as a "go-to" provider of trusted solutions. Many services are on offer, including: dashboards hosted inside Learning Management Systems (LMSs); Student Information Systems (SISs); student risk models; Learning Record Stores (LRSs); adaptive learning systems; and analytics products that provide insights for decision makers about planning, resource allocation and best practice pedagogies. As a result, institutions are starting to differentiate EdTech products on the basis of their analytics capabilities.

However, it is important to recognise that educational institutions maintain a *portfolio* of technology products and services to deliver learning experiences, and that integrating data solutions across this portfolio remains challenging. It is often claimed that common standards across these products and services will ensure that the data generated in the resulting learning ecosystems can be integrated in an interoperable format (Stringer, DeMonner and Heyer, 2017), and so standards-based approaches are increasingly recognised as important components of mature, production-level learning products. The two most common standards for learner activity data are IMS Caliper and the Experience API (xAPI).

The central questions that this paper tackles are: ***how useful for understanding learning processes is the data resulting from the products institutions purchase? And are there productive ways to resolve limitations as they are identified?***

The paper is organised so that different stakeholders can dip in and out according to their expertise and interest. Section 1 provides an overview for the nontechnical reader, introducing the problem of generating data for learning analytics ecosystems, and exploring why many institutions are not currently getting the data they need to improve learning outcomes for their students. In Section 2 we make a series of recommendations, aimed at three different sets of stakeholders: (i) institutions procuring educational technology; (ii) vendors making that technology; and (iii) the researchers and practitioners in the LA and EDM community. Section 3 consists of a set of technical appendices that provide specific examples of problems that we have encountered, along with evidence for our claims in the earlier sections.

For the sake of clarity, it is worth emphasising that we believe the current sub-optimal situation can be resolved only in partnership with technology vendors, and that most are good actors seeking to make the best product possible with the resources that they have available. This paper is intended to provide all stakeholders with a deeper understanding of the problems currently faced by those building learning analytics solutions. We hope that this paper will assist researchers, practitioners, and educational institutions in working together to realize deeper educational insights, and the resultantly better outcomes for our learners, that we all aspire to.

A number of different concepts will be referred to throughout this position paper, many of which are used interchangeably by various stakeholders who use educational data. We will make use of the following definitions in what follows;

- **Institutional Analytics:** the use of data, statistical analysis, exploratory and predictive models to gain insight and act on complex issues related to business practices (see Brooks and Thayer, 2016)
- **Web Analytics:** the measurement, collection, analysis and reporting of web log data for purposes of understanding and optimizing web usage
- **Learning Analytics (LA):** 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 (see Lang, Siemens, Wise and Gasevic, 2017).
- **Educational Data Mining (EDM):** the development of methods for exploring the unique types of data that come from an educational context (see Romero, Ventura, Pechenizkiy and Baker, 2010).

Section 1: The Problem

It is important to begin by considering how the clickstream data used in LA was initially created - it owes its historical origin to the logs maintained by technology applications (i.e. *telemetry data*). These were initially created to assist with debugging technical problems — not with understanding the behaviors of students and other users. Making the connection between telemetry data and educationally relevant analytics has turned out to be a challenge that is often overlooked. If products claim to supply educationally-relevant LA, but only provide technically-oriented data and analytics, the results will struggle to provide useful insights into teaching and learning.

Data data everywhere... but where are the insights?

In many cases, significant investments are being made in educational data ecosystems without an rigorous evaluation of whether the data collected can actually be used for intended (or planned future) purposes. This has resulted in an increasing number of practitioners and researchers discovering that it is difficult (if not impossible) to use the data so far collected, sometimes at great expense, to make educationally valid inferences about students, educators, and the learning environments which they are using. This position paper stems from a recognition among LA researchers and practitioners that the data being generated by vendors and stored in institutional databases is largely unfit for its stated purpose. Many in the community are watching with growing unease as large sums of money and enormous effort are devoted towards the production and collection of data that we worry will not provide the types of deep educational insights that LA researchers aim to provide, and that educational institutions increasingly expect. A risk is emerging that key stakeholders responsible for purchasing EdTech solutions will start to doubt the value of educational data and de-prioritise it in their acquisition processes.

Generating better educational data requires a long term strategy. One that moves beyond short term convenience, and towards a systematic solution that provides relevant stakeholders with the data they need to help learners, instructors, and the institutions which aim to support them.

Will customers keep investing in data and analytics?

As people who consume educational data, often working with solution providers to improve its quality, we have made a number of mistakes, and learned a number of lessons. We hope that sharing our experiences will help the growing number of new entrants to the LA/EDM space to avoid (some of) the mistakes that we (and others) have made during early attempts to build LA ecosystems that have data feeds with the potential to help us understand and support learning processes.

Purpose-built design leads to data that provides deeper educational insights

The most important lesson we have learned is that educational data must be developed using a sound theoretical framing about what it is that we are trying to measure, rather than trying to align existing clickstream and telemetry data with constructs after-the-fact.[1] Best results are obtained if educationally meaningful constructs are identified first, and then used to inform both what we are measuring (i.e. the relevant fields) and the protocols that we use to measure it. Starting with low level clickstream and telemetry data tends to result in gaps and oversights within the data generated, and often it is impossible to repurpose a dataset collected opportunistically towards educationally meaningful use cases. For example, a common clickstream activity is "view" or "play" without reference to the context of the object being "viewed" or "played". And yet, the difference between a student viewing a quiz item and a learning resource are dramatic, and must be separated if we are to gain meaningful insights into student learning. This is no different from any other technology application; the intended user value should drive the creation of the technical solution that is created, not the data stream.

End user scenarios must be considered *first.*

Our current educational data infrastructure is not usually developed with this grounding. Instead, software developers drive the development of standards and data formats without the input of product managers or other experts with knowledge of customer data requirements. This means that we often find ourselves in a position where the *wrong* data is emphasised, and great care taken to ensure that this data is provided in a robust technical format (see Appendix D for an example of this phenomenon). We do not dispute the necessity of using robust standards to generate and collect data, but proportional focus should be given to ensure that the right data is being created in the first place. After all, the most interoperable data in the world will rapidly become useless if it fails to support any relevant use cases. In Appendix H we argue that it is likely to be easier to generate interoperable data if priority is given to the production of data that maps easily to its anticipated end uses. From here we can map back from theoretically well-founded cognitive, social and educational constructs; to the analytical methods that will be used with the data; to the data required to deliver this analytics; and finally to the required syntax and semantics of that data required. Unfortunately, this process is rarely followed in producing educational data.

1 See also the recent blog post by Michael Feldstein at https://eliterate.us/pedagogical-intent-and-designing-for-inquiry/ for a further example of the need to link learning intent to data capture.

But why are things going wrong?

In what follows we will briefly highlight some of the underlying issues that have led to this state of affairs, before moving onto a discussion of how we think the EdTech industry could move forward to achieve better outcomes. We start by examining the core issues that we believe are driving current practice.

A lack of attention to learning ecosystems

Educational data science is at a point of transition, with focus turning from analytics for data collected by individual tools and towards analysis of student behaviour across *learning ecosystems*, consisting of a wide range of tools and EdTech products. More and more institutions are attempting to merge data from the many different sources and learning tools in their ecosystem including: large scale LMS and SIS solutions; small interactive tools that are integrated with the LMS via mechanisms such as the Learning Tools Interoperability (LTI) standard to improve the learner experience; and specialised solutions that support curriculum design, student feedback, language proficiency assessments, and more. Many institutions have discovered that data from one platform rarely accounts for the breadth of the student experience. For example, a student with low LMS usage but high eText usage may be highly engaged in a course, but a report produced using LMS data alone would fail to account for this.

As institutions grapple with the increasing complexity of modern learning, we have seen concepts like the Next Generation Digital Learning Environment (NGDLE)[2] emerge, which seeks to integrate multiple tools using a component based architecture to deliver the learning environment of the future. Most educational institutions make use of a wide range of teaching technologies, and need the EdTech sector to work towards facilitating data analysis across all of the resulting data streams. This is because many of the educationally relevant constructs that concern us (e.g the status of a student as being “at risk” of failure, or their collaboration with team members in a group project) require the integration of data from more than one stream. Beyond the campus supported learning ecosystem, instructors make use of wordpress sites, clickers, and their own favoured online tools. Similarly, our students make use of social media, youtube videos, MOOCs and other resources on the web to enhance their learning experience. Even beyond the K-16 education sector, employers are working to help their employees learn on the job and re-skill as entire industries are disrupted by AI, automation, IoT and robotics. The range of EdTech that a person will interface with over their lifetime of learning is vast.

2 https://er.educause.edu/toc/educause-review-print-edition-volume-52-number-4-july-august-2017

Learning Ecosystems: A learning ecosystem consists of people, designed learning activities, and the tools that support them. These tools often include a collection of EdTech tools and products.

For LA to achieve its intended value within this context, we must work to make use of data from these disparate environments, and many more that have yet to even be invented. It is unreasonable to expect that the needs of increasingly data savvy institutions can be addressed by analytics that are provided by a single application. We frequently witness a move towards extracting data from stand alone systems so that it can be linked across multiple applications in more sophisticated analytics solutions capable of delivering richer insights. As of today, there are few examples of successfully building multi-application insights, although some initial work has been conducted (Bakharia et al., 2016; Forteza et al., 2017, Kitto, et al., 2020). It is becoming apparent that the education sector needs vendors to move beyond attempts to constrain analysis within a single standalone system, and towards enabling genuine data interoperability across applications. Institutions that use multiple systems will not settle for anything less.

A lack of consistent data policies and practices

Who at an institution is empowered to represent its entire data and analytics needs? We commonly see competing or conflicting requirements emerging, often from a lack of coherent policies and procedures. Learning Analytics can be developed in faculties as a research project, in central units that are in charge of technology and its support, in learning and teaching units that are charged with improving the student experience but have no control over the EdTech systems that they use, and more recently in hybrid models that seek to combine the best aspects of all of these (Buckingham Shum and McKay, 2018). Consistent policies are needed across an institution if it is to develop robust and educationally relevant learning ecosystems. Indeed, as Reinitz (2019) claims: “Effective analytics is a cross-enterprise effort that requires effective data governance.” All institutional stakeholders should be provided with avenues for influencing decisions that will affect them. A data service that fails to support all of these groups and their individual (sometimes conflicting) requirements leads to a solution that supports none of them. We have all seen situations where an EdTech product is purchased with insufficient consideration of how it might affect instructional affordances, or promising LA projects which have been shut down because of a shift in LMS, institutional policies, or other disjointed decisions made with little awareness of the broader institutional learning ecosystem.

A clear vision for LA (Fergusson, Brasher et al., 2016) helps to align this complex set of interdependencies. However, the visions of different educational institutions will not always be the same, which creates challenges for vendors if they seek to adopt a "one size fits all" approach. Who owns the data? Who can access? Will students have a right to erasure? If so then what data can be erased? How long is data stored for? Is there an opt out option? For everything? Or only some types of data? All of these questions (and more) must be answered by each institution as it starts to link up data into an ecosystem that collects data about individuals across multiple systems and activities.

Consistent policies are also needed between vendors, institutions, and the people being described by the data that LA seeks to use (i.e. our students). A distrust is emerging in the education sector about big tech companies and their questionable privacy policies. These must be directly addressed by both the companies tracking student behaviour in their systems, and the institutions that are seeking to use this data to improve student outcomes. There are a number of learning analytics initiatives that have begun to address this issue, such as the SHEILA framework from the EU, which was developed and validated with over 200 participants (https://sheilaproject.eu/). Much more work remains to be completed in this area, but the LA community is notable for its ongoing interest in this area. See for example special issues on the topics of privacy (Ferguson, Hoel, Scheffel, and Drachsler, 2016), the ethical use of student data (Buckingham Shum and Luckin, 2019), and the need for participatory design of LA solutions (Buckingham Shum, Ferguson, Martinez-Maldonado, 2019). The LA community has also created a rich vein of more technically orientated discussion papers (Corrin et al., 2019), reports (Drachsler et al., 2016), codes of practice (Sclater & Bailey, 2015) and policies implemented by institutions acknowledged as leading the field[3] which all serve to demonstrate a keen interest in treating learners' data with the utmost respect. Rogue actors that fail to do so are not only failing to use best practice, they bring the entire field into disrepute.

Poor implementations of standards

Data standards like Caliper and xAPI are often touted as the way in which we can move towards interoperable data that is usable across a learning ecosystem. The claim is frequently made that the quality of our educational data will improve if we make use of standards. For example, the IMS Global Learning Consortium claims that the Caliper standard "Provides a standard way of measuring learning activities and effectiveness, which will enable designers and providers of curriculum to measure, compare and improve quality"[4], but in Appendix C we present a scenario where the data captured using this very same standard has resulted in no such improvement. Nevertheless, claims that

3 See e.g. https://help.open.ac.uk/documents/policies/ethical-use-of-student-data and https://www.ed.ac.uk/academic-services/policies-regulations/learning-and-assessment/learning-analytics

4 See e.g. https://www.imsglobal.org/initiative/real-time-cross-application-educational-data-and-analytics (accessed Sunday 8th March 2020).

we just need to follow data standards are by no means rare, and imply that one of the reasons for poor quality data comes from problematic (or non-existent) implementation of data standards. Why would this occur?

Standards exist for a reason, and should be followed, contributed to and extended wherever possible. However, short term thinking often drives the build of a technical solution - time, cost and the experience of a team all factor into the delivery of a product that is worthwhile, and it is rare to see organisations invest the requisite sums of money. This leads to shortcuts, pragmatic decisions, and a lack of interest in scanning for best practice examples that could inform the current product cycle. It is often challenging for businesses to prioritize standards-driven development of data instrumentation given that it does not immediately result in a product feature that can be released. As Stringer, DeMonner and Heyer (2017) have noted, there are numerous *social problems* facing standards adoption, including: the slowness of the standards development process; researcher resistance; diversity in tools; institutional issues; and a lack of vendor cooperation. How might we start to do better?

Even when standards are followed by an EdTech solution, there is an emerging realisation that the resulting data often fails to usefully inform educational decision making (see Appendices A-D). The people designing standards are commonly of a more technical background, with less understanding of educational theory or practice, which often means that the resulting data is not readily mapped into educationally relevant concepts. Providing data that can support learning is not as straightforward as predicting flows through a website, or interventions that will maximise sales. The criteria for success are far more difficult to evaluate (Gašević, Dawson, S. and Siemens, 2015). In many cases the data currently being produced is too fine-grained over educationally irrelevant concepts, while being too coarse-grained over concepts that are of key importance for the field (see in particular Appendix C).

Of course, the problem of implementation has not been helped by the fact that there are a number of *competing* standards.

Competing standards

One of the problems that has received significant attention in recent years[5] is that of competing standards. Two activity data standards have emerged in the EdTech community, Experience API (xAPI, https://github.com/adlnet/xAPI-Spec) and IMS Caliper (https://github.com/IMSGlobal/caliper-spec). While the code bases for both are open source, each has a markedly different set of core principles underlying the standard, development model and

5 See e.g. the Edinburgh statement on data interoperability, released during LAK'16: https://github.com/AlanMarkBerg/hack-at-lack16/blob/master/TheEdinburghStatement-Signed.pdf

adoption pattern. Consolidating the standards is difficult for both cultural and technical reasons. At this point in time we believe that both standards will remain in use across various EdTech platforms for the foreseeable future. This situation has caused considerable confusion for both educational institutions and EdTech vendors who would seek to support them. Large vendors usually support at least one specification (e.g. Instructure Canvas now has Caliper certification, and Articulate Storyline supports xAPI), and sometimes both (e.g. Blackboard is certified for Caliper and is a recognised xAPI adopter). We are aware of vendors being approached by different institutions with varying demands to satisfy one or the other specification to support their internal data models, a situation that adds to the development costs of a product (if the onus is put on the vendor) or which shifts the load to institutions (if they need to map data emitted natively from one product into their implemented data models). This is not an optimal state of affairs.

Institutions making new decisions around learning products and data analysis platforms should review the standards that are supported by these products and take this factor into account when making purchasing decisions. It is not clear at present whether either standard has a better result for educational institutions at the standard level. In Appendix A we discuss the way in which this confusion impacts upon the relatively simple scenario of collecting data about video events. Not only do we have a problem with poor data being provided, Appendix A shows us that even when a data stream adheres to a standard it rarely provides all of the information that is required by LA practitioners. As institutions receive data from multiple providers in different formats, we see institutions repeatedly being forced to make significant data processing and cleaning investments. This situation will increasingly frustrate practitioners and lead to cynicism about the utility of data standards.

Attempts to treat learning data as a unique competitive advantage hurt us all

As the strategic value of data and analytics becomes more widely understood in EdTech it is no surprise that a wide number of interests (including those of individuals, institutions, companies, collectives, and consortia) are all leveraging what strengths they have to make market gains. Companies that provide more useful and relevant data want to realize a financial advantage for this achievement, just as any other superior product feature should convey a market reward. However, many of the questions asked by decision makers require the combination of data streams from various products (e.g. many LMS solutions now use the LTI standard to plug in new products created by other vendors and their associated extra functionality). This makes it important to create data in an open, standards-compliant fashion that facilitates reuse in other systems and products. Boutique custom approaches to data may provide a short-term market advantage for a vendor creating a product (e.g. an excellent "in solution" dashboard can be pitched as a competitive advantage during sales pitches), but the value of data in the broader ecosystem is dramatically reduced if the data cannot be extracted and combined with the data collected in other systems. There is a danger that decision makers will start to question the value of the data and analytics products that they purchase if these fail to answer pertinent questions about student learning across the entire learning ecosystem - and an associated decline in the perceived value even of best practice data streams and analytics tools.

Short term thinking will harm the data & analytics sector

In Appendix F we discuss a scenario that is starting to occur where vendors recognise that simply following the REQUIRED aspects of the Caliper specification tends to result in data that is not particularly useful for LA. They are starting to implement OPTIONAL features of the specification on their own, responding to immediate use cases, to provide better analytics for those purchasing 'their' product. A similar scenario has occurred in xAPI where not all vendors are adhering to the best practice Profile specification, resulting in a wide range of data formats just for the simple case of video (see Appendix A) which has flow on consequences already: some LRSs actively market their solution as working particularly well with these non standard but popularly used solutions. These ad hoc responses to market need are reasonable short-term strategies, but they have the potential to restrict the educational insights that can be gained from the LA ecosystem in the long term.

Data is rarely the topmost priority when an institution purchases EdTech products (although it is increasingly becoming an important one), which means that requirements elicitations will normally list data format, its accessibility and interoperability at a "nice to have" level of importance. Core to an acquisition process will be a business need - user experience and integration with existing systems are justifiably often given a higher priority. This means that most large institutions are now in a position where they are dealing with data coming from multiple sources, some of which will follow a standard, but many of which do not. It is increasingly becoming necessary to ingest and then analyse data that adheres to multiple types (e.g. CSV, JSON, JSON-LD, and XML), and standards (e.g. xAPI, Caliper, PAPI, Activity Streams), regardless of the competitive drivers of those who produce it. The best data governance strategies in the world are struggling to devise methods that simplify the process of turning this data into something that can yield actionable insights. Learning analytics ecosystems require whole system thinking from each stakeholder - we need to incentivise a shift towards products that create data that can be used across the ecosystem.

How can we start to break this vicious cycle emerging from our current practices? In what follows we sketch out recommendations for three classes of actor in the system: institutions procuring EdTech products; EdTech vendors; and practitioners in the LA/EDM community.

Section 2: Recommendations for addressing the problem

Recommendations for institutions procuring EdTech products

Vendors are starting to face situations where different institutions and consortia are making differing demands about adherence to data standards. Not only is it possible that vendors can be asked to deliver both xAPI and Caliper, different data quality requirements will increasingly start to enter the process of tendering for products. This will cause frustration on both sides if we do not work towards harmonising data policies and expectations. This process should be conducted with reference to the robust literature that has been developed by LA/EDM researchers over the past decade. This section proposes a number of steps that can be taken by institutions to ensure that the data they receive from products will be more likely to meet their needs.

1. Start with a clear vision and applied project for analytics.

Clarity of purpose is key to achieving learner and educator engagement with analytics projects. As is discussed by Ferguson et al. (2019), there are many possible visions about what purposes data and analytics might serve at your institution, and it is important that you clearly define and articulate this vision in advance of making any product decisions. For example, a solution that monitors the classroom to support teaching and learning requires very different data, and very different privacy controls, from one that seeks to return personal data to students to help them create evidence of learning gains over a lifetime. Conversations about potential analytics projects should set the first deliverable as a set of educationally-relevant questions that the proposed project could help answer. Providing a clear articulation of what problem you are trying to solve using data (which may well change in time) will help those involved in the tender process to produce use cases and requirements that will fit with a strategic, cohesive data strategy. This vision will also help to identify the value and priority of analytics vs. other requirements.

2. Ensure that your LA ecosystem is supported by a robust and justifiable ethical framework.

A failure to develop a sound ethical framework for your LA ecosystem could threaten its very existence[6]. The complex socio-cultural environments in which learning occurs require a nuanced approach to this difficult challenge - a one size fits all approach is unlikely to work. However, numerous examples of sound ethical frameworks that have been

6 See e.g. the InBloom controversy and the Stchiting Snappet case that are discussed in Drachsler et al. (2016).

implemented at leading institutions can now be found, and can serve as a source of inspiration for those seeking to develop a LA ecosystem.[7]

3. Require that raw data (event-level, streaming) be made available outside the application.

Don't assume that data and analytics provided by a single product will prove adequate in the long term. Even if a dashboard or LA tool meets your current needs, it is likely that you will very soon start asking questions about what that data means when it is combined with other datasets in your learning ecosystem. This process will be facilitated if all of the tools in your system make use of a common data format and vocabulary that exposes data at the "event-level" (e.g. "Jamie opened the syllabus."), rather than in an aggregated format (e.g. "Jamie opened six files."). It will also prove much easier to implement if best practice APIs and modular architectures are provided as standard components of the product. See Appendix E for details about the types of information that we often find missing in data streams provided by products we have utilised. Requiring such data during negotiations with vendors would facilitate the linking up of datasets to give a more comprehensive model of student behaviour across the entire learning ecosystem.

4. Include analytics experts in the core team to draft tenders and review responses.

Analytics solutions may include technical components that require vetting and review by people with experience in statistics and data mining, ideally within the domain of education. These experts should be considered core members of the tender committee to help identify and select solutions that can meet the requirements of the stated project vision. Including such people in an acquisition process can ensure that your organization considers data and analytics at a deeper level than simply ticking off a box, and that robust analytical methods are in fact being used by the product to be acquired. This is especially critical in the current market, where we now have many solutions delivering "analytics" at varying levels of maturity. Some of these use robust approaches, but many others do not.

5. Require best practice implementation of standards.

Tender documents should specify data standards that your institution prefers. Furthermore, they should be explicit about which Metric Profiles (for Caliper) or Profiles and Recipes (in xAPI) are expected. This can help to ensure that the data generated across multiple products is interoperable, thereby saving substantial investment post acquisition

7 Ethical Use of Student Data for Learning Analytics, The Open University: https://help.open.ac.uk/documents/policies/ethical-use-of-student-data
Learning Analytics Principles and Purposes, University of Edinburgh: https://www.ed.ac.uk/files/atoms/files/learninganalyticsprinciples.pdf

in tasks like e.g. resolving data types (e.g. IRIs) that cannot be displayed in user facing tools, data manipulation, and the mapping variables (e.g. different user IDs for one user) when attempting to join datasets. One way to ensure this best practice in the products that you acquire is to develop an active engagement with standards bodies, working to ensure that the practices embodied in those standards are aligned with the lived experience of your institution.

6. Inspect a data sample to ensure useful data is provided.

We have all too frequently experienced scenarios where a solution has promised an API architecture, and “all the data” only to find out post purchase that its capabilities were far more limited than we expected. This is not from malice - different vendors and institutions have very different expectations as to what is required for analytics, but it is reasonable to expect that your institution will increasingly start to require more data, at a finer granularity, and for a wider set of purposes than originally expected. Inspecting a sample of the data provided by a new product, and ideally trying to use it to satisfy a well defined institutional need, will quickly clear up many misconceptions about whether the data is indeed fit for your intended purpose.

7. Form a bloc to develop an agenda and reduce required effort.

Some of the institutions that have achieved the best successes with their LA infrastructure are those that have formed a bloc and negotiated with vendors as a collective. Jisc’s Effective Learning Analytics project[8], and the Unizin consortium[9] both provide an example of this model. In each case a bloc of universities are represented en masse, which enables both vendor and institutional certainty. This approach can help to generate more robust and detailed requirements than a single institution can usually achieve alone.

Recommendations for EdTech Vendors to generate useful data

What can vendors and those delivering EdTech products do to ensure that their data and analytics solutions are as useful as possible for their customers? How can they build data streams and solutions that become increasingly valuable as more analytics work is conducted? Generating data that will be useful to institutional business owners and end users requires a deeper understanding of how data will be used and toward which goals. It is not enough to simply gain a stamp of certification or to comply with a standard if that doesn’t provide useful data to solve authentic problems. Ongoing collaborations with educational clients and LA/EDM experts can provide insights into the actual

8 https://www.jisc.ac.uk/rd/projects/effective-learning-analytics
9 https://unizin.org/

needs of educators, but deep relationships like this are built over a long time period and take investment from both sides. We believe that many of the currently adopted approaches to product delivery are short-term in scope and should start to include a longer-term orientation to ensure ongoing customer satisfaction.

8. Educational scenarios should drive data generation, not technology functionality.

We must move from a situation where internal telemetry and simple descriptions of product features are driving the generation of the data formats provided to those purchasing a product. This approach is inherited from other sectors which focused on mining "data exhaust" for unexpected patterns using statistics or machine learning, but in education, this inevitably leads to a situation where a multitude of clickstream data cannot be mapped to educationally relevant constructs (see Appendices C and D), which creates a disconnect when it comes to providing educators and learners with meaningful feedback. It is essential that vendors and the data standards community start to consider the broader use cases for this data. A top down approach (see Appendix H) *designs data* from educational constructs. This has long been understood in measurement/assessment science (e.g. Milligan and Griffin, 2016; Mislevy, Behrens, Dicerbo, and Levy, 2012), and the integration of such approaches into LA has been shown to yield higher quality data than simply instrumenting a set of functionalities that originate in tracking users through websites (see Appendix D, or e.g. Shibani, Knight and Buckingham Shum, 2019; Wise, Knight, Buckingham Shum, In Press). Learning is a more complex process than e.g. web browsing, with deeper potential insights if properly understood and instrumented.

9. Follow data standards at the best practice level.

Following data standards will help to ensure that institutions do not have to constantly "reinvent the wheel" in order to utilise the data they obtain in the various solutions they procure across their learning ecosystem. However, this will only work if vendors adopt the best practice implementation of those standards, viewing the data as something that must flow to other products and solutions, rather than be trapped within "their" system. Common and open vocabularies must be developed and used across the *sector*. For example, simply adopting xAPI does not entail best practice adoption (see appendix A) - it requires the generation and use of profiles which are stored on the server. Similarly, much work could be completed to align vocabularies across xAPI and Caliper, and an openly created resource linking these two specifications should be prioritised as a genuinely collaborative endeavour (i.e. all stakeholders must be involved in the process, not just the members of specific communities).

10. Implement more than REQUIRED.

All standards bodies tend to produce specifications that categorise data types at the REQUIRED, RECOMMENDED, PREFERRED, OPTIONAL, etc. levels. This results in many vendors choosing a minimal adoption model where they implement just the REQUIRED levels and ignore the remainder of the specification. While it is certainly possible to claim conformance with a specification through such a model in the short term, this leads to poor outcomes in the long term as much of the data most useful for LA/EDM has only an optional status in current standards. Many vendors are currently implementing both xAPI and Caliper statements at a minimal level, and failing to deliver best practice data according to the standard as it is currently written (see Appendices A, B and C) which further decreases the utility of the resulting data. In Appendices E and H we discuss how the entire sector might work towards improving this situation.

11. Pay attention to your competitors' data formats and work towards interoperability with them.

Enabling a rich lifelong learning ecosystem that facilitates the many and varying needs of institutions requires that all designers and builders of data and analytics solutions work together to achieve genuine data interoperability. The time is rapidly approaching when a single data stream from one product will no longer be considered acceptable. This means that vendors must work together to ensure that the data generated by their solutions works seamlessly with other vendor's products. A short term competitive advantage arising from a proprietary "better data service" that is built from non-standard uses of data has the potential to hurt all members of the sector in the long term, as key decision makers fail to see a return from investing in a data and analytics infrastructure that cannot be used more widely beyond the purpose for which it was initially built. It will also make integration with increasingly complex institutional learning ecosystems more difficult, as a lack of coordination means that institutions will start to demand that "their" data policy be adopted. This wastes time and effort across the sector. While much has been learned in trying out various data formats and technologies, as the field matures it is becoming important that we work towards harmonisation across our learning ecosystems - we should be consolidating the lessons we have already learned across the industry. There are many mechanisms for achieving a collaborative working relationship with institutions, other vendors, and even competitors. They include SoLAR SIGs, working parties, the ISO and IEEE. We encourage the entire sector to concentrate on building up strong communities of practice and defined mechanisms for information transfer between all stakeholders.

12. Prioritize the long term.

It will not be possible to maintain a competitive advantage on the basis of the data that you "hold" for much longer. More and more institutions are demanding that data generated by students in their learning ecosystems is returned to them. Will they be able to fruitfully use this data to support learners in the full complexity of their learning process at this point? Not if the data returned follows a unique format, or is a poor implementation of an existing standard. Is your product going to support this complex process? Institutions will only continue to invest in data and analytics products if they deliver a worthwhile return on investment, beyond the immediate requirements of an acquisition. Care must be taken to ensure that educational data is free to meaningfully flow, across product based, temporal, sectoral and geographical boundaries, as this will both reduce the barriers to investment (because analytics can be reused across various products and geographical domains). Vendors would do well to compete on the richness of their data, its interoperability with products supplied by other vendors, and the transparency of the analytics they provide. Successful vendors will provide the best quality data service which supports institutions as they work to support learners not just in learning concepts in the syllabus, but also in developing higher order capabilities such as critical reflective thinking, creativity, excellent skills in collaboration, and communication.

Recommendations for the LA/EDM community

The LA/EDM community itself must also take some responsibility for improving the current status quo. LA/EDM practitioners are starting to be given long sought access to institutional data stores, only to discover that the data collected is insufficient to address the questions that they have been tasked with answering. A general reluctance to participate in the standards making process means that we are now starting to find ourselves confronted by data which was touted as "interoperable" but which does not serve our needs. How can we work to rectify this state of affairs?

13. Get involved.

The LA/EDM community will only receive high quality data if it helps to create it. Stringer, DeMonner and Heyer (2017) claim that researcher resistance is a key factor in resistance to the adoption of standards, pointing to a lack of engagement in the community. If experts in the learning sciences leave the process of standards development to experts in technical development then we will only create ongoing problems for ourselves in the future. The time to get involved is now.

14. Doing your own thing is easier at first... but hurts us all over time.

As Feldstein (2017) points out, specifications tend to produce compromises that few people are thrilled with. Usually it is easier to do your own thing when building a new method or prototype. However, this form of pragmatism hurts us all over time, and is likely to lead to long term cynicism among institutional decision makers about the cost vs benefits of LA/EDM. If all data and analytics tools make use of different syntax, semantics, and data structures then the costs of implementation will be too high. To evolve niche research prototypes towards a maturity level where they can be institutionally adopted we must find ways to work together with vendors, standards bodies and the wider community to deliver tools and methods that make use of common standards, and interoperable data formats.

15. Think about how to scale your research from the outset.

Many interesting research avenues have emerged from LA/EDM over the decades, but few have been implemented at scale as trusted products in institutions. Many of them cannot be used beyond the classroom where they were first implemented by a researcher, others fail to take into account important pedagogical aspects of a class and adapt accordingly. For LA/EDM to mature we must work as a community to build models that can achieve sustainable impact (Buckingham Shum and McKay, 2018). This process starts with those performing research - while initial results can be interesting in their own right, a mature solution only emerges by considering all factors associated with implementation and scaling (Knight, Gibson and Shibani, 2020).

16. Help others to follow in your wake.

Open data and reproducible research are fast becoming an accepted norm for scientific research.[10] And yet EdTech is littered with projects that were funded for a period of time, and then dropped as the researchers involved moved onto new projects. We often see innovations that have to be completely reimplemented from scratch, because the original codebase is inaccessible, cannot be updated/understood, and the person originally responsible for it can no longer be found. Available documentation with strong implementation guidance is needed to help those who would follow in your wake to understand how a tool or resource works, what it communicates, and how it should be used. Open software repositories, published datasets, technical HOWTOs and other methods to enable longevity should all be considered throughout the life of a project.

17. We need a scientific community driving standards development.

The standards development community started with the challenge of engineering solutions that provide baseline functionality, such as authentication between systems, creating rosters of enrolled students, etc. These are important

10 https://osf.io/

processes, however, they are not the same as delivering useful data and analytics, a process which requires many careful considerations (e.g. semantics, feature selection and the higher level problems associated with algorithms and sensemaking). The scientific community (e.g. LAK/EDM) is often somewhat removed from the standards development process, perceiving it as dull, and slow. And yet the standards development process affects us all. We need to work towards developing an active LA/EDM community that helps to guide the standards development process, without requiring the direct involvement of data science researchers and practitioners in the process itself. Instead of accepting what we are given, the LA/EDM community must seek to shape the data infrastructures that we will increasingly have to work with. Developing new models for influencing this process must become a strategic priority of all communities involved in the learning sciences. One model for this can be found in the recent announcement of a collaborative effort to develop a lesson level interoperability standard[11], but more need to be developed. Many other avenues are possible, big and small, such as participation in Special Interest Groups (SIGs), writing position papers that push best practice forwards, encouraging students and technical staff to develop prototype tools using best practice from the outset. As a research community we need to encourage a collective research agenda that includes implementation and replication studies, not just new algorithms and analyses. What steps can you take to ensure that we get useful data?

In conclusion

While there is currently interest and investment in learning data ecosystems, many learning analytics researchers and educational data scientists are encountering challenges when attempting to answer educationally meaningful questions. We have argued that this is often due to the low quality data that they are forced to work with. If this situation is not addressed, it is likely that institutional decision makers will start to de-prioritise the acquisition of data infrastructure, analytics applications, and research efforts - thereby hurting all stakeholders in the data and analytics community. This is a situation that can be fixed, but only if all stakeholders step up to the challenge.

This paper has identified a number of political, technical and commercial reasons why the data in our learning ecosystems is currently failing to support student learning as well as it might. We have also offered recommendations directed at various stakeholders which would help to improve the quality of our educational data. Better quality data, collected in an ethical manner, that helps us to answer identified educationally relevant questions, would provide learning analytics practitioners with a rich resource that would help us to provide our students with a better learner experience, more useful feedback, and rich personalisation. Such data, collected from well designed learning ecosystems, would provide institutions with a critical enabling resource that will help us to improve the outcomes of the many and varied learners that we cater for, at all stages in their unique lifetimes of learning.

11 https://eliterate.us/announcing-a-lesson-level-interoperability-standards-effort/

Section 3: The Evidence (Appendices)

These appendices provide examples and more detailed discussion to support the arguments raised in the main discussion above. They are intended for a technical audience.

A: The already messy case of video events

To illustrate just how wrong things can go for those building learning analytics ecosystems we can look at the seemingly relatively simple scenario of video events. It is quite likely that anyone building LA infrastructure at a large educational institution (e.g. a university) will be confronted by a variety of data syntax, just for the case of video. To be more specific, we currently have:

1. An xAPI video profile[12] which specifies a common vocabulary and statement structure that SHOULD be followed when implementing xAPI video events.
2. A Caliper media metric profile[13] which will increasingly be adopted by major providers.
3. Kaltura following a slightly different xAPI statement structure[14] despite the published video profile.

Few institutions make use of just one video player. Most use YouTube or Vimeo LTI integrations with their LMS, interactive eBooks with their own players, and other stand alone tools with a video component. This state of affairs means that LA/EDM practitioners are starting to receive data in all of the above formats, as well as other data formats that come from vendors which don't follow any standard. When an instructor or administrator asks to see statistics about video usage at the organisation we are commonly left with partial data, as some data is not retrievable, and often cannot report on all video usage at an institution. The LTI connections provided by many LMSs provide a good example of this problem as they do not pull data generated by the tool back to the referring tool. Such data must be collected separately, which is not always possible.

Video is a relatively simple use case, with well understood events, and a set of associated questions that educators would commonly ask of any LA tool which served their purposes. For example:

12 https://github.com/adlnet/xapi-authored-profiles/tree/master/video
13 https://github.com/IMSGlobal/caliper-spec/blob/master/caliper-spec.md#36-media-profile
14 https://watershedlrs.zendesk.com/hc/en-us/articles/360022749332-How-do-I-author-video-xAPI-statements-

- Did the student who just asked me how to do the assignment previously watch the video describing the assignment?
- Which videos are being watched most frequently in my course?
- Did students watch an entire video? Or just a snippet?
- Which of my videos get the most views?
- Are there parts of my videos that get a lot of views? Why?
- What were students doing before they watched a video I provided?
- Where did a student watching a video I provided go next?

The simplicity of the video use case helps us to demonstrate some of the issues that arise in attempting to answer these fairly common questions. In what follows we will unpack the types of problem that a data analyst is likely to encounter in the current state of affairs generated by a wildly disparate data ecosystem. Consider for example the following problems that a data analyst frequently encounters.

Problem 1: It is hard to compare usage between different video tools

This problem arises very quickly in a LA ecosystem. Consider an authentic example of the type of problem that will be increasingly faced by LA practitioners in a large institution which uses two video players: Kaltura and a fictional BestPractice product which has implemented the xAPI Profile for video.

Consider Figure 1, which contains two xAPI example statements[15] concerning a video event which *should* be largely equivalent. They both track a user viewing a video, and both are legal xAPI statements. However, both store markedly different information. Even at the level of verbs, one statement uses `viewed`, and the other `completed`, and it will be necessary for a LA practitioner to map between those two verbs for all video events at the back end before they can start to compare usage patterns between the two tools. Note that only the *Bestpractice* tool emits a timestamp, despite this being essential information for LA practitioners who often want to explore usage patterns over a time window. The *Kaltura* statement provides no information about how long the video was, although there are other examples on the same page for `played` and `watched` events, which both include percentage amounts for *how far* through the video the user has gone. However, this percentage form of information is not the raw data that LA

15 Accessed from https://watershedlrs.zendesk.com/hc/en-us/articles/360022749332-How-do-I-author-video-xAPI-statements on 19th July, 2019.

practitioners tend to prefer when attempting to map data across systems, and a further question presents at this point - are the three Kaltura events listed on the source webpage largely the same? Why are they using different verbs if so? While both statements have implemented some of the information necessary for LA in `context`, neither of these statements include the text that could be used in a dashboard providing a user readable description of what class or activity the video was embedded in (necessitating further data collection to resolve the URI), and both of them provide different data in this field. More issues that are likely to arise when mapping between these two markedly similar events can be found by browsing through the two code samples. Thus, a quick examination of the data highlights a number of problems that would arise for an institution in asking even very basic video usage questions.

Problem 2: It is likely that the practitioner can't identify specific users across tools

Some of the most common questions that a practitioner might want to ask concern correlations between video usage and student participation in other activities. For example, we often want to determine whether students who watched a video performed better on an assessment task, or to correlate pausing and rewinding through different concepts in a video with quiz scores, performance in a lab task and on the job performance. While it is common to want to join datasets in this way, this is a nontrivial task in a learning ecosystem. Different providers will often make use of different identifiers for a user, and it is left to the institution to map between them. Even if two tools have instrumented Caliper there is no guarantee that user identifiers will be resolvable between statements generated by them (see Appendix E). Identity management, a crucial aspect of user modelling, remains a highly challenging problem.

Problem 3: Information essential to the learning sciences is not captured

Finally, even with the best practice adoption of either xAPI or Caliper, key information is missing from both specifications. For example, no attention has been paid to the learning design in which a video appears, despite strong results emerging in recent years to suggest that this has a profound effect upon student activity traces (Rienties and Toetenel, 2016). Similarly, information about the intention that a video is meant to serve is not stored. Is it part of an extension activity? Or is it explaining how to complete an assessment task? These two intents should lead to vastly different student behaviour in watching a video, and yet essential contextual information of this type is currently missing from all data specifications. This is due to an emphasis upon low level clickstream data, and an associated ignorance about what data is necessary to achieve a positive impact upon student learning. A common standard that could be used to tag such constructs might start to alleviate this problem, but this would require a significant investment from the standards, vendors and researcher community, as well as a significant extension to what is currently considered as best practice.

xAPI Video Profile – Completed


```
{
    "actor": {
        "mbox": "mailto:video.user@example.com",
        "name": "Video User",
        "objectType": "Agent"
    },
    "verb": {
        "id": "http://adlnet.gov/expapi/verbs/completed",
        "display": {
            "en-US": "completed"
        }
    },
    "object": {
        "id": "https://examplevideosite.com/media/01898390",
        "definition": {
            "type": "http://adlnet.gov/expapi/activities/video",
            "name": {
                "en-US": "An introduction to corporate financial
reporting metrics"
            },
            "description": {
                "en-US": "An introduction from our head of finance
into the metrics we report on as a business."
            }
        },
        "id":
"http://video.examplecdn.net/v/corporatefinaicialreportingmetrics.mp
4",
        "objectType": "Activity"
    },
    "result": {
        "completion": true,
        "extensions": {
            "http://id.tincanapi.com/extension/ending-point":
"T1H3M55S"
        }
    },
    "context": {
        "contextActivities": {
            "category": [
              {
                 "id": "http://adlnet.gov/expapi/activities/video"
             ]
        },
        "extensions": {
            "https://w3id.org/xapi/video/extensions/session-id":
"7a1f8a80-8b62-4cf5-9165-a4976280f6c4",
            "http://id.tincanapi.com/extension/duration": "T1H3M55S"
        },
        "registration": "3da7f040-2a1b-11e9-b210-d663bd873d93"
    },
    "timestamp": "2019-02-06T14:05:25.158Z",
    "id": "40db117a-2a1b-11e9-b210-d663bd873d93"
}
```


Kaltura – Viewed


```
{
  "actor": {
          "mbox": "mailto:video.user@example.com",
          "name": "Video User",
          "objectType": "Agent"
  },
  "verb": {
    "id": "http://id.tincanapi.com/verb/viewed",
    "display": {
      "en": "viewed"
    }
  },
  "object": {
    "objectType": "Activity",
    "id":
"https://example.mediaspace.kaltura.com/default/channels/view/channe
lid/75565771",
    "definition": {
      "name": {
        "en": "Example Channel"
      },
      "type": "http://id.tincanapi.com/activitytype/category"
    }
  },
  "context": {
    "contextActivities": {
      "category": [
        {
          "objectType": "Activity",
          "id": "http://tincan-
definitions.kaltura.com/products/mediaspace/1",
          "definition": {
            "type": "http://id.tincanapi.com/activitytype/source"
          }
        }
      ]
    }
  },
  "version": "1.0.0",
  "id": "fbea8f92-3983-4649-8134-4ef8eb44fd94"
}
```


Figure 1. Two example statements about video usage, produced by two different tools: *BestPractice*, a fictional tool that has adopted the xAPI Profile for video, and *Kaltura*, who have implemented their own xAPI statement structure.

B: Minimal adoption models

Despite the considerable amount of work that has been completed in the standards community for both xAPI and Caliper, the requirements for obtaining status as an 'adopter' are still very weak. For example, in the IMS Caliper scenario, a tool need only instrument one metric profile before a vendor can enter into the IMS certification process for that profile, and if the tool passes all tests then it will be badged as IMS Caliper certified. However, this does not mean that *all* Caliper metric profiles relevant for this tool will be implemented. An institution that trusts the badge without exploring what profiles are supported by a tool may be bitterly disappointed about what knowledge can be drawn from the data it provides. Similarly, it is possible for a vendor to generate xAPI statements in their tool, and thereby claim xAPI compliance on their webpages and documentation, but, as we have seen from Appendix A, there is every reason to expect that xAPI statements coming from two different vendors will be markedly different in their semantics, and that this will cause significant problems when attempting to use the resulting data.

This problem tends to arise because far too many fields which we consider essential to LA are declared as OPTIONAL in both specifications. In xAPI for example, the optional **`context`** property has a further set of minimal properties (**`registration, instructor, team, contextActivities, revision, platform, language, statement`**, and **`extension`**) all of which are also optional. While some of these have obvious values, the **`contextActivities`** and **`extension`** properties have proven to be particularly difficult for those building LA infrastructure. Different applications will use **`contextActivities`** in a range of different ways, and extension is frequently used as a catch-all, with wildly varying structure.

While reducing required components is a sensible decision to make in terms of minimising the cost of adoption, it carries a risk that those investing in data and analytics infrastructure will not be able to address even basic questions with the data they manage to collect, despite making significant investments, and will become disappointed with the state of the technology and eventually halt those investments.

C: Caliper / xAPI event streams that say almost nothing

Caliper metric profiles and optional fields provide a vocabulary with the ***potential*** to describe a broad array of learning events. Similarly, xAPI allows an organization to create a recipe that makes detailed descriptions of activities. However, in practice the implementations of these two specifications often falls short, providing data streams with largely redundant events of a single type, or include system-generated events that fail to describe any user behavior. This scenario is not addressed through e.g. Caliper certification tests, as these tests only evaluate if the system is able to send a properly-formatted event using the Caliper protocol. The fidelity of a given data event from Caliper (or xAPI) to a user activity that ***should*** generate that event is not evaluated - and would be difficult to satisfy through an automated test, as evaluating fidelity requires making statements around intent and meaning of a learning activity.

The result of this poor implementation is a data stream that may produce a large number of "events" (i.e. millions of records), yet provide almost no insights or other value to an academic institution. An example of 10M Caliper records generated by two weeks of LMS activity is provided in Figure 2. This table is from a real world implementation of a learning ecosystem at a large institution: 76.3% of the records describe a "view" event (with no contextual information about what type of object was viewed provided within the stream alone) and 11.5% describe a system activity; less than 10% of the events remain. Very few useful insights can be generated with this data.

This situation is not unique to Caliper; we have seen many xAPI data streams generated over the years which are large in volume but low in variety. Without a diverse range of behaviors represented, this data cannot be used for educationally-relevant learning insights. Several of our recommendations above address this issue; from institutions requesting a data sample to vendors ensuring that their instrumentation is directed by use cases. We note that Caliper statements include a way to provide related learning objectives,[16] we would like to see more methods like this developed as a priority. Longer term, initiatives such as the lesson level interoperability effort recently launched[17] as a part of the Empirical Educator Project (EEP) will help to ensure that clickstream data is coupled to intent, learning design, and educationally meaningful constructs.

Viewed	7,843,347	76.3%
Graded	1,179,002	11.5%
Started	327,913	3.2%
Completed	327,885	3.2%
LoggedIn	194,649	1.9%
TimedOut	180,278	1.8%
LoggedIn	66,443	0.6%

Figure 2: Redundancy in 10M Caliper action records

16 https://www.imsglobal.org/sites/default/files/caliper/v1p1/caliper-spec-v1p1/caliper-spec-v1p1.html#learningObjective
17 https://eliterate.us/announcing-a-lesson-level-interoperability-standards-effort/

D: Web traces vs learning events

It is likely that one of the causes behind the poor data currently being generated in learning technologies stems from an assumption that LA is just web analytics applied to learning. This is not terribly surprising. Wikipedia describes web analytics as “the measurement, collection, analysis and reporting of web data for purposes of understandingandoptimizing web usage”[18], a definition that shares many features with the SoLAR definition of LA. However, optimizing web usage is not the same as “optimizing learning and the environments in which it occurs”[19], and when vendors conflate the two this leads to the generation of data that is not particularly useful for LA.

Many vendors use web analytics to establish key performance indicators (KPIs) as quantifiable measures of success. In our observation of current vendor practices, stakeholders for a product’s success tend to define such KPIs by the measures of what they can readily measure and understand:

- Monthly Active Users
- Usage (broken down to a given user)
- Traffic (broken down by how the product is accessed).

A vendor will also look at how active the accounts are, and who in the accounts is accessing their products. Such KPIs are *product*-centric, i.e. generally collected by vendors to help improve the product. Often this telemetry data is what gets operationalised in APIs and specifications about learning data, but such data is only a part of what the field of LA requires.

The Learning Sciences are a well established field (Sawyer, 2005), and much has been learned over the past decades about what makes learning different from other scientific endeavours. In education we are rarely seeking to optimise click-through rates, or to maximise sales (although these are no doubt sometimes priorities for our websites). We are far more concerned with identifying things like: what instructional methods are most effective; the best practice approaches to teaching students metacognition and reflection; identifying students who are in trouble; and extending those who are currently doing well. Simple aggregations of web activity will very rarely support this type of understanding. However, people continually enter the domain of IT support for EdTech from sectors like banking, web development, and marketing, and they bring with them inappropriate assumptions about what education needs. Old mistakes, which many in the field thought were dealt with and gone, frequently emerge and cause considerable

18 Web analytics. (2019, May 21). Retrieved May 21, 2019, from https://en.wikipedia.org/wiki/Web_analytics
19 Retrieved 4th June, 2019 from https://en.wikipedia.org/wiki/Learning_analytics

consternation and wasted effort. Consider for example the apparently never-ending attempt to adjust products to student “learning styles” despite decades of evidence suggesting that this fails to improve learning in any way whatsoever. An excellent introduction to what psychological constructs actually *do* correlate to academic performance can be found in Richardson, Abraham and Bond (2012); learning styles do not. Some recommendations for the type of data LA actually requires are given next.

E: What basic information should be in a useful statement?

There are a number of fields that we consider essential to building LA solutions, which are not currently being delivered by many EdTech learning data streams:

- **Object Names/Titles:** Objects should include human interpretable names e.g. if an object was a course then a course title (e.g. "Psychology 101") should be included. A failure to include this information puts a heavy onus on LA solution builders, who must either work to resolve an IRI to text, or display information in dashboards etc. that is not comprehensible to humans and therefore unlikely to be useful in a vast range of contexts.
- **Object IRI's:** Object extensions should provide resolvable links to the relevant resource e.g. https://jisc.instructure.com/courses/3. This is not mandatory in the xAPI spec, but is necessary for resolving data to use it in deeper analysis. Continuing with our example above, resolving the "Psychology 101" subject would ideally also enable a LA solution builder to find an IRI to the relevant curriculum information, as this would enable them to analyse activity against curriculum changes, and perhaps even work towards the automated recognition of prior learning. We note that IMS CASE[20] identifiers provide a potential method for including interoperable information about what competencies are being taught via an Object, and so offer a potential method for resolving Object IRIs with more useful information.
- **Object information:** Information such as instructor details, learning objectives, teaching methodology, learning design etc. are increasingly becoming important in the most sophisticated approaches to LA. And yet this information is rarely stored in a form that can be accessed.
- **Object membership:** Clear identification of object membership is necessary for much of the reporting that institutions require. For LA practitioners it is important to know that an item belongs to a course/module so that activities can be tracked on that basis. For example, although Canvas Live Events data does provide a 'membership' object which defines the relationships between content, this is often not populated, e.g. data about wiki pages which belong to a group or course is often not returned with that information provided.
- **Object External IDs:** Where objects represent a concept that exists outside the system (e.g. course, student) they should include a meaningful ID rather than just an internal systems id. For example, using 'PSY101' to represent 'Psychology 101' is likely to result in information that is not useful in the long term. Similarly, if a Student ID is not resolvable beyond the system in which the data was generated then it will be impossible to link data across the LA ecosystem.

20 https://www.imsglobal.org/activity/case

As we recommend above, when acquiring new tools and solutions which will form part of an institution's LA ecosystem it can be extremely informative to inspect data samples ahead of time. During this process we recommend that LA practitioners be asked to build quick displays using the sample data, attempting to use it to display insights of relevance to the institution.

F: The Dangers of Extending Events Ad Hoc

One approach that vendors have taken to dealing with missing data in standards and specifications is to extend event profiles in an *ad hoc* manner to include the items that are useful in the case of their specific application. This is a rational approach for vendors to take - it helps them to provide meaningful insights as quickly as possible. However, without bringing these changes back to influence the required specification, negative downstream implications arise, including:

- **Lack of consistency:** if each vendor takes an ad hoc approach, key attributes will be defined differently; not only in what they are called (e.g. "logout" vs. "end session") but the way that information is represented could be different (e.g. time of logout vs. duration of session).
- **Application-specific data mapping and interpretation:** This lack of consistency leads to substantial additional work to decode each event stream that is incorporated. It can mean that each event stream must be interpreted separately at the time of integration into a data warehouse (or research project). This is the same situation that is encountered if no standard is used. The end result is more work for both the vendor to create a standards conformant event and no savings to the application user/consumer, a lose-lose situation.
- **Loss of effective practices and research outcomes:** there will be better and worse ways to represent learning events. Research will uncover attributes that are found to be more and less useful at describing meaningful learning processes. These findings will only be used for other applications if a consistent approach to events is used. By contrast, if a consistent approach was used, findings from one study or analysis could have broad implications across the field.

Resolving this situation requires a closer relationship between the technical teams implementing standards and those responsible for defining and expanding the specification. The lessons and adaptations that are being made to extend default events should be brought back to influence enhancements and revisions to the events.

G: Ensuring interoperability between xAPI and Caliper

Performing an exact mapping between xAPI and Caliper is unlikely to be feasible given the current state of the two specifications. While some claims have been made that convergence is achievable, it is not straightforward to map in either direction between the two standards as they are currently formulated:

- **Mapping Caliper to xAPI:** Much of Caliper must be mapped to the `extensions` field inside the `result` and `context` fields of an xAPI statement, which loses essential semantic information. Extensions can literally be anything, which makes it almost impossible to generalise analytical approaches over them. Thus, a comprehensive mapping of Caliper to xAPI results in many statements which, while syntactically correct, are not particularly useful to LA/EDM. In the current format, an xAPI based analytics architecture would struggle to do anything interesting with statements that were originally generated in Caliper and then converted.
- **Mapping xAPI to Caliper:** Conversely, many xAPI statements have yet to be represented as Caliper metric profiles. This means that there is no way to map them at all. However, as Caliper extends its domain of application there will be more and more points where convergence should be sought (see Appendix A which discusses the case of video events in the two specifications).

But a complete map is unlikely to be necessary

However, things are not as bad as they might seem, partly because most implementations of both xAPI and Caliper currently adopt a fairly minimal approach (as noted in Appendix B). We believe that a useful starting point will be to review the *actual* use of xAPI and Caliper used in production systems today, rather than trying to map between the entire specifications. Convergence can focus on commonly implemented elements, with new, shared developments focusing on constructs that are meaningful for learning analytics. We will discuss why this is an approach likely to yield useful results in what follows.

H: Top down vs bottom up data capture

LA practitioners are increasingly starting to recognise that the best way in which to generate useful data is to start by considering use cases and work backwards to data capture. Thus, rather than working from the telemetry data that is captured through web traces, standardising it, and then starting to think about the analytics and reporting that could be created from it, we believe that better data is generated from a "top down" approach that is informed by best practice research in the learning sciences (Knight, Buckingham Shum, 2017, Kitto, Buckingham Shum et al. 2019). This state of affairs has arisen more from a historical artefact than bad intentions. Tool creators have tended to start with basic usage metrics as a first step, and we are now finding that these map poorly into educational outcomes despite their simplicity of implementation. Such an approach leads to the "clicks to constructs" problem (Whitmer, San Pedro, et al., 2019), where it is almost impossible to map low level clickstream data into educationally relevant higher order constructs which are useful for analysts and senior decision makers. Now, as the industry matures we must start to adopt a more science informed approach where we seek to measure what matters. We can do this by working from the top down, and thinking at the outset in terms of educational constructs and the potential end uses of the data being collected leads to better data. This approach requires extensive collaboration between educational experts and the tool creators who would seek to support them.

Educational constructs help to ensure data interoperability

Interestingly, such an approach is also likely to provide a way forward as we seek to resolve the thorny problem of mapping between the two specifications without losing important information. Consider for example the case of writing analytics (WA), which is becoming an important substream of LA. It would be easy for a solution provider who is building a tool to support student writing to instrument data collection from the web traces perspective (see appendix D). In such an instance a tool would provide data about: what page in a LMS a student opened up the tool from; how long a student used the tool for; how many words they write in that session etc. This information is sometimes useful (depending on the aim of an analysis), but WA tends to focus upon a very different set of features, such as: the rhetorical moves a student makes to support an argument (Gibson et al., 2017); the linguistic features they have in their writing (McNamara, Graesser, McCarthy, and Cai, 2014) etc. While many of these features can be extracted from raw text using off the shelf tools, we have seen few EdTech providers emitting either xAPI or Caliper data that includes such text. The variable most interoperable across tools (raw text) tends to be missing.

A similar lesson was learned during the BeyondLMS project[21], which sought to map data collected from a variety of social media environments into educationally relevant LA using xAPI. Developing the Connected Learning recipe (Bakharia, Kitto et al., 2016) led to far more useful data than would have been achieved if a simple instrumentation of social media data had been carried out. It also ensured that data was mapped interoperably between a wide variety of social media tools - a situation which would have been unlikely to occur had the tools been integrated one at a time, or on an *ad hoc* basis.

21 www.beyondlms.org

References

Bakharia, A., Kitto, K., Pardo, A., Gašević, D., & Dawson, S. (2016). Recipe for success: lessons learnt from using xAPI within the connected learning analytics toolkit. *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 378–382). ACM. https://doi.org/10.1145/2883851.2883882

Brooks, D. Christopher, and Terri-Lynn B. Thayer. Institutional Analytics in Higher Education. Research report. Louisville, CO: ECAR, March 2016. https://library.educause.edu/-/media/files/library/2016/2/ers1504ia.pdf

Buckingham Shum, S., & McKay, T. A. (2018). Architecting for learning analytics: Innovating for sustainable impact. *Educause Review*, 53(2), 25-37. https://er.educause.edu/articles/2018/3/architecting-for-learning-analytics-innovating-for-sustainable-impact

Buckingham Shum, S.J. & Luckin, R. (2019). Learning Analytics and AI: Politics, Pedagogy and Practices. *British Journal of Educational Technology*, 50(6), pp.2785-2793. https://doi.org/10.1111/bjet.12880

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-Centred Learning Analytics. *Journal of Learning Analytics* (Special Issue Editorial), 6(2), 1–9. https://doi.org/10.18608/jla.2019.62.1

Corrin, L., Kennedy, G., French, S., Buckingham Shum S., Kitto, K., Pardo, A., West, D., Mirriahi, N., & Colvin, C. (2019). The Ethics of Learning Analytics in Australian Higher Education. A Discussion Paper.https://melbourne-cshe.unimelb.edu.au/research/research-projects/edutech/the-ethical-use-of-learning-analytics

Ferguson, R., Brasher, A., Clow, D., Griffiths, D., & Drachsler, H. (2016). Learning analytics: visions of the future. LACE Technical Report, Workpackage D3.2, http://laceproject.eu/wp-content/uploads/2016/02/LACE_D3_2.pdf

Ferguson, R., Hoel, T., Scheffel, M., & Drachsler, H. (2016). Guest Editorial: Ethics and Privacy in Learning Analytics. Journal of Learning Analytics, 3(1), 5-15. https://doi.org/10.18608/jla.2016.31.2

Feldstein, M. (2017). What is the next generation? *EDUCAUSE* Review:38–44, 2017. Available at: https://er.educause.edu/articles/2017/7/what-is-the-next-generation

Forteza, D., Whitmer, J., Fritz, J., & Green, D. (2018). *Improving Student Risk Predictions: Assessing the Impact of Learning Data Sources*. Blackboard. https://www.blackboard.com/resources/improving-student-risk-predictions-assessing-impact-learning-data-sources

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, 59(1), 64-71.

Gibson, A., Aitken, A., Sándor, Á., Buckingham Shum, S., Tsingos-Lucas, C., & Knight, S. (2017). Reflective writing analytics for actionable feedback. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 153-162). ACM. http://dx.doi.org/10.1145/3027385.3027436

Griffiths, D., Hoel, T., & Cooper, A. (2016). Learning analytics interoperability: Requirements, specifications and adoption.Public Deliverable D7.4. LACE Project (European Commission Seventh Framework Programme, grant number 619424). Retrieved from http://www.laceproject.eu/d7-4-learning-analytics-interoperability-requirements-specifications-and-adoption/

Drachsler, H., Greller, W., Griffiths, D., Hoel, T. Kickmeier-Rust, M. (2016). Is Privacy a Show-stopper for Learning Analytics? A Review of Current Issues and their Solutions. *Learning Analytics Review*, no. 6, January 2016, ISSN: 2057-7494. http://www.laceproject.eu/learning-analytics-review/privacy-show-stopper/

Kitto, K., & Knight, S. (2019). Practical Ethics for Learning Analytics. *British Journal of Educational Technology*, 50(6), pp.2855-2870. https://doi.org/10.1111/bjet.12868

Kitto, K., Lupton, M. Bruza, P., Mallett, D., Banks, J., Dawson, S., Gašević, D., Buckingham Shum, S., Pardo, A., Siemens, G. (2020). Learning analytics beyond the LMS : enabling connected learning via open source analytics in 'the wild' : final report. Dept of Education, Skills and Employment, Available at: https://ltr.edu.au/resources/ID14-3821_Kitto_Report_2020.pdf

Knight, S., & Buckingham Shum, S. (2017). Theory and learning analytics. *The Handbook of Learning Analytics*, (Chapter 1), pp.17-22. https://doi.org/10.18608/hla17.001

Knight, S., Gibson, A., & Shibani, A. (2020). Implementing learning analytics for learning impact: Taking tools to task. *The Internet and Higher Education*, Vol.45 (Apr. 2020). https://doi.org/10.1016/j.iheduc.2020.100729

Lang, C., Siemens, G., Wise, A., & Gasevic, D. (Eds.). (2017). *Handbook of learning analytics*. SOLAR, Society for Learning Analytics and Research.

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Milligan, S. and Griffin, P. (2016). Understanding learning and learning design in MOOCs: A measurement-based interpretation. Journal of Learning Analytics, 3(2), 88– 115. http://dx.doi.org/10.18608/jla.2016.32.5

Mislevy, R. J., Behrens, J. T., Dicerbo, K. E., & Levy, R. (2012). Design and discovery in educational assessment: Evidence-centered design, psychometrics, and Educational Data Mining. Journal of Educational Data Mining, 4(1), 11–48. Retrieved from http://www.educationaldatamining.org/JEDM/images/articles/vol4/issue1/MislevyEtAl V ol4Issue1P11_48.pdf

Reinitz, B. (2019). Why Data and Analytics Matter—and Why They Matter Now: Report from the EDUCAUSE/NACUBO/AIR 2019 Enterprise Summit. https://er.educause.edu/blogs/2019/8/why-data-and-analytics-matter-and-why-they-matter-now

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological bulletin*, 138(2), 353.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior*, 60, 333-341. https://doi.org/10.1016/j.chb.2016.02.074

Romero, C., Ventura, S., Pechenizkiy, M., & Baker, R. S. (Eds.). (2010). *Handbook of educational data mining*. CRC press.

Sawyer, R. K. (Ed.). (2005). *The Cambridge handbook of the learning sciences*. Cambridge University Press.

Sclater, P. & Bailey, P. (2015). Code of practice for learning analytics. Jisc. Available at: https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics

Shibani, A., Knight S. and Buckingham Shum, S. (2019). Contextualizable Learning Analytics Design: A Generic Model, and Writing Analytics Evaluations. *Proceedings 9th International Conference on Learning Analytics & Knowledge*, Tempe, AZ. https://doi.org/10.1145/3303772.3303785. Available at: https://tinyurl.com/lak19clad

Stringer, J., DeMonner, & Heyer, O. (2017). The promise of learning data interoperability. *EDUCAUSE* Review, 2017. Available at: https://er.educause.edu/articles/2017/7/the-promise-of-learning-data-interoperability

Whitmer, J., Liu, R., Walton, K. E., Moore, J. L., & Lotero, A. A. (2019). *The Constructs Behind the Click*s. Technical report, ACT, Inc. https://www.act.org/content/dam/act/unsecured/documents/R1803-la-sel-2019-11.pdf

Wise, A., Knight, S., Buckingham Shum, S. (In Press) Collaborative Learning Analytics. In: Cress, U., Rosé C., Wise, A. & Oshima, J. (Eds.), *International Handbook of Computer-Supported Collaborative Learning* (Springer). Preprint available at: https://bit.ly/3lXtsFZ

Authors

Kirsty Kitto is an Associate Professor in the Connected Intelligence Centre (CIC) at the University of Technology Sydney (UTS). Her research straddles complex systems modelling, cognitive science, and learning analytics. At UTS:CIC she works to create learner facing tools that help people learn how to learn over a lifetime of transitions. An early adopter of xAPI in the LA community, she is known for her work on educational data interoperability and portability, with a goal to facilitate student ownership and control of their learning data, over a lifetime of interaction with many different learning ecosystems.

John Whitmer, Ed.D. John is a Senior Fellow for the Federation for American Scientists to the Institute for Education Sciences, a position which allows him to apply his decade of data science experience to innovative projects to reduce costs, improve learning environments and student outcomes. Prior to this role, John led teams of data scientists, research scientists, and machine learning engineers in large educational assessment companies (ACTNext) and edTech providers (Blackboard). These teams worked on a variety of foundational research and production projects, such as automated writing assessment using natural language processing, predictive analytics using clickstream and background data, adaptive testing platforms, and business intelligence applications. An educational researcher by training, John approaches these projects with a commitment to improving the lives of underserved and marginalized students. John holds a Doctorate in Educational Leadership from UC Davis and a Master's Degree in Sociocultural Anthropology from UC Davis.

Aaron E. Silvers consults on learning analytics strategies to drive and improve practical, real world outcomes. Through content examples, tutorials and open source projects related to data analytics, adult learning and virtual collaboration, Aaron is known for his early output of tutorials, documentation and SCORM eLearning content examples that demonstrated how multimedia-based eLearning could communicate with LMSs, enabling the advent of eLearning authoring tools. Aaron is also recognized for his stewardship of the open-source xAPI community and his roles in the creation of both xAPI and xAPI Profiles. Aaron maintains an active role internationally in the development of tools, technologies, tutorials and standards benefitting governments, healthcare, energy, higher education, and other organizations that develop and depend on people. In addition to his role with Elsevier Inc, architecting a modern learning analytics capability, Aaron consults government, military and concerns for learning data at scale through MakingBetter (http://makingbetter.us/). In Fall 2020, with the release of a first xAPI Profile Server designed and developed for ADL, Aaron chairs the IEEE P9274.2 xAPI Profiles Standardization effort. aaron@makingbetter.us.

Michael Webb is director of technology and analytics at Jisc. Jisc are the UK higher, further education and skills sectors' not-for-profit organisation for digital services and solutions. He led the development of Jisc's learning analytics service - a national service for higher education in the UK, and is currently working on a range of innovation projects using AI to support learning, teaching and research. Before joining Jisc he worked in a number of higher education institutions in the UK, providing IT services and solutions.

Acknowledgements

We are indebted to our expert review panel for their feedback on earlier drafts of this Position Paper.

Alan Berg, University of Amsterdam

Michael Feldstein, e-Literate

Linda Feng, Unicon Inc.

Jason Paluck, University of Maryland, Baltimore County

Etienne Pelaprat, Unizin

Anthony Whyte, University of Michigan

SOLAR

SOCIETY for LEARNING ANALYTICS RESEARCH

September 2020