An Integrated Approach to Learning Analytics

Uchendu Nwachukwu[1], Sandy Ingram[1], Juan Carlos Farah[2], and Denis Gillet[2]

[1] College of Engineering and Architecture of Fribourg, Switzerland
{uchendu.nwachukwu,sandy.ingram}@hefr.ch
[2] École Polytechnique Fédérale de Lausanne, Switzerland
{juancarlos.farah,denis.gillet}@epfl.ch

Abstract. Despite the rise in services generating learning analytics, there is a lack of standard models and guidelines for data integration and aggregation to inform the design choices of applications supporting learning analytics. We propose a bottom-up, user-driven approach enabling educators to select, match, and contextualize activity traces from several data sources to perform and visualize meaningful learning analytics. To facilitate the process, the proposed approach recommends building customized auxiliary plugins that can be shared and re-purposed. We present the implementation of a use case following this approach. This use case focuses on supporting the import and side-by-side comparison of activity traces from multiple data sources that teachers might use in their practice. Implications of this approach on cross-platform learning analytics and future work are discussed.

Keywords: Digital Education, Learning Analytics, Interoperability, Data Integration

1 Introduction

Studies on the use of learning analytics (LA) have garnered substantial interest over the last decade [16]. This interest has been further emphasized by the sharp shift of learning activities to online settings as a consequence of the COVID-19 pandemic [8, 13]. LA are considered promising resources to support learning [25]. Although LA are often generated within Learning Management Systems (LMSs), online learning activities often comprise interactions with services outside the traditional LMS (e.g., video streaming services, scientific databases, interactive applications) [7]. When several applications are used for the same course, it becomes important to consider activity traces coming from different sources in order to enable an overall assessment of the learning activities [6, 18].

Over the past decade, there has been a strong push for the adoption of specifications defining the storage format of activity traces. One of the leading formats is the Experience API (xAPI), which has been acknowledged as a standard and is currently used by over 200 systems [17]. Nevertheless, there are challenges for interoperable LA, even with established standards. Systems might apply different vocabularies for the same actions, most notably demonstrated by the example of social media platforms (e.g., retweeting vs. sharing a post) [11]. Furthermore,

some systems rely on other specifications, such as IMS Caliper [10] or Activity Streams [26]. Thus, it is not straightforward to merge data from multiple data sources. Given that educators are often the ones selecting the technology they adopt to support their practice [15], we believe that they are in the best position to understand the semantics of each data source used and determine the activity traces that are significant considering the learning context at hand.

To address this gap, we propose a bottom-up approach to assemble activity traces from various sources, on-demand. More specifically, our approach delegates to educators the choice of selecting what input data to map, analyze, and represent. We simplify this process by sharing auxiliary plugins and thus enabling the reusability and extensibility of solutions. The possibility of sharing plugins adds a social or crowdsourcing dimension to our approach, allowing instructors to collaborate on solutions that might be applicable to multiple learning scenarios. Furthermore, our approach proposes a transparent integration of static and dynamic LA reports. These reports allow students to reflect on their learning process, while also enabling teachers to adapt their teaching methods even while a course is still ongoing. To exemplify our approach, we present a use case involving a sample course using two separate learning platforms; namely Moodle [9] as a repository for learning material and quizzes, and Graasp [5] as a collaborative space with dynamic content, interactive applications, and virtual labs. In this use case, a proof of concept shows how activity traces coming from one of those platforms are fed and consumed by the other, producing LA that can be then embedded into the corresponding learning space.

2 Related Work

Students and teachers form the primary target audience for LA [21, 28]. Feedback for teachers is usually provided at the course level [22], with the aim of improving their teaching methods [16]. While existing studies mainly focus on the analysis of single-data-sourced LA [19], combining multiple data sources gives better coverage and overview of learning activities [16].

LA platforms can be loosely or tightly coupled with the LMS where the corresponding course activities take place [4]. If bound to an LMS, LA are performed through native functionalities or plugins, where data is accessed directly from within the LMS [3]. For LMS-independent solutions, there exist various open LA frameworks [14]. In such frameworks, data is gathered in a central learning record store (LRS) that might store activity traces from multiple learning tools sharing a common terminology standard such as xAPI or IMS Caliper [14].

Several LA metrics can be found in LMSs and in the literature. For instance, the Moodle Engagement Analytics Plugin measures the *student engagement* metric using indicators such as *logins*, *forum participation*, and *assignments* [3]. A study from 2017 [23] used the above-described metric to predict student course outcomes but could not find any significant correlation. However, it did so when considering only the *login* indicator. This study, among others, sheds light on a core challenge for LA: the lack of valid and standardized metrics that are widely

adopted and interpretable. The choice of LA metrics highly depends on the learning context [1, 20]. A course may take place in a blended learning environment or purely online, just as the content and learning targets of the course may vary. Hence, a metric may be suited for a particular context but not another [12, 24, 27].

Educators are key actors in determining relevant LA metrics [2, 29]. This paper takes into account the lack of standardized LA metrics and the strong dependence of LA metrics on their usage context, thus proposing an educator-centric, standard-independent, peer-to-peer development and communication approach to collecting and exploiting LA inputs, outputs, and tools. To start with, by distinguishing separate roles in the process of producing LA, the proposed approach recommends the adoption of a component-based architecture, easing the reuse, composition, and re-purposing of LA tools. Furthermore, the approach calls for peer-to-peer distributed communication between multiple data sources through connectors, without the need for a central LRS. The result is a procedure that supports cross-platform LA without the need for a standard to be adopted by the learning platforms themselves. In the following section, we present this procedure in detail, focusing on the role different components play in enabling our integrated approach to LA.

3 Integrated Approach to LA

In the absence of widely adopted common standards for LA metrics and for combining activity traces from heterogeneous sources, we propose a high-level, bottom-up, educator-driven approach where activity traces are fetched on-demand from various data sources, analyzed, and exploited following a component-based and peer-to-peer architecture, as represented in Figure 1. The core characteristics that guide our approach are described in this section.

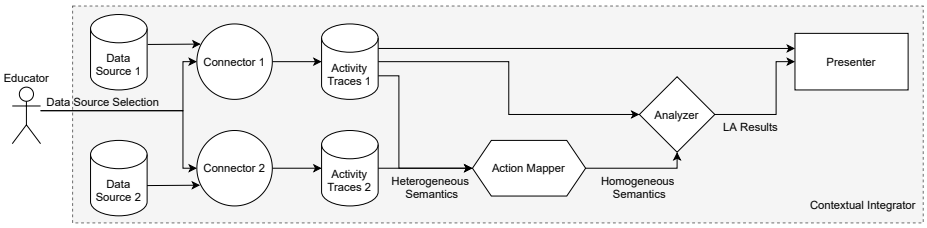

Fig. 1. The user-driven approach lets the educator select connectors to import relevant activity traces from data sources and optionally create a mapping for the vocabulary used. This mapping enables the combination of activity traces that might be using heterogeneous semantics. The combined data source is then used as input for an analyzer, which then passes the output to a presenter that makes it consumable by the educator. Alternatively, the presenter can be directly coupled to the combined data source to explore the available data.

3.1 Educator-Centered Development

As noted in Section 2, educators are often the ones behind the choice and design of the learning scenarios underpinning LA. Educators are also best suited to find and interpret meaningful information that might be captured by LA, and—consequently—to communicate that information to students, parents, colleagues, and other stakeholders, which vary according to the context in which the learning is taking place. Therefore, we propose taking a context-specific and educator-centric approach—i.e, where the educator plays a central role in the process of designing, re-purposing, configuring, and publishing data collection, analysis, and visualization tools. In practical terms, the educator is granted full end-to-end control to decide which activity traces are interesting, where to store them, and how to make use of them. Furthermore, the educator interpolates verbs coming from different data sources and selects an appropriate representation for the output. This interpolation is indeed a first step in supporting consistent cross-platform LA, in which activities that happen in one platform can be linked with or mapped to those that happen in another.

3.2 Shareable and Composable Organic Enablers

Our approach encourages the organic development of shareable enablers, which consist of loosely coupled components that can be used individually or together based on the context at hand. Built independently by end-users, they each cover a distinct aspect of the LA process: (1) connectors fetch activity traces (2) mappers specify and map actions of interest, (3) analyzers perform advanced data analysis (e.g., anonymization, descriptive or prescriptive analysis), and (4) presenters report the outcomes to end-users. The communication between different enablers is neither standardized nor predefined; it is rather directly delegated to them. Furthermore, when multiple data sources are involved, data combination is not enforced at any specific level; it can either be done at the analysis level or only at the presentation level. This lack of centralization and standardization is another key step in supporting cross-platform LA. Platforms are not required to be in concert with each other or adapt their data storage or communication formats. Instead, different platforms—and integrations between them—can become supported by enablers developed by the end-users of such platforms.

In addition to the four different enabler types—which are listed below—our approach identifies two key components: (1) contextual integrators and (2) publishers. Contextual integrators represent the learning environment or space integrating enablers and their outputs. These integrators provide situated and contextualized LA storage and access solutions without depending on a third-party external LRS. Publishers, on the other hand, expose enablers to educators to facilitate discovery and re-use, and can serve as open repositories of LA solutions. These repositories can be curated by experts or crowdsourced by the community.

Connectors fetch learning traces such as activity logs, quiz results, or course completion states from data sources. For every source of interest, a dedicated connector is required to import data. The connector should be developed in an educator-driven way. Typically, it must include filtering features facilitating the selection of a subset of activity traces of interest.

Action Mappers enforce the same semantics to define actions described by activity traces when using multiple data sources. Therefore, an action mapper should let the educator match synonyms to unify the data. The output of such a component is ideally a reusable configuration that can be shared and extended.

Analyzers grant descriptive or prescriptive insights into the available data. Analytic options are handled separately from the import logic because the same analytic option might be applied in different data contexts. This makes them reusable and encourages collaboration and sharing through an open community.

Presenters provide feedback in the form of visual representations or textual reporting. The input can be raw activity traces coming from a connector or elaborated data coming from an analyzer. In a bottom-up approach, the available data determines the output format. If there is no suitable format available, it can be developed and afterward shared with the community. Presenters exploiting different data sources can be put side-by-side in a dashboard. This side-by-side grouping helps the user get a better overall assessment of the output.

3.3 Contextual Integrators

The approach favors the integration of the LA process in the learning environment, where relevant activity traces from different sources can be combined and stored within the context of a course. Typically, both enablers and their outputs (e.g., raw activity traces, LA, dashboards) should be easily added to the online learning space or environment of a course. Integrating enablers and outputs into the learning space paves the way for just-in-time analytics that can also be useful for students taking the course.

3.4 Publishers

The separation of concerns and the on-demand development of enablers represent the core characteristics of the proposed approach. It requires educators to put more effort into designing their LA scenarios and develop (or ask for the development) of enablers. At the same time, our approach also envisions the process of collecting and analyzing learning traces as an open and collaborative one. As enablers including connectors, presenters, mappers, and analyzers—for more elaborate usage scenarios—are developed and shared, and the community around shared enablers grows, the task of developing enablers shifts to searching

for the right enablers, while designing LA scenarios shifts to exploring scenarios that have been reported as useful in similar contexts. Publishers expose enablers to educators by relying on high-level descriptions without enforcing a rigid representation or exchange standard. Essentially, publishers provide repositories of enablers that can be grouped according to the specific needs of the publisher's audience (e.g., subject matter, language, curriculum).

4 Use Case Implementation

In this section, we present a proof-of-concept implementation of a use case based on our proposed approach. The use case scenario is first described and the state of its implementation follows. It is worth noting that the work presented in this paper focuses on the data import aspect of the connector and provides basic presenter functionalities. Ongoing and future work will be discussed in the final section.

4.1 Use Case Scenario

It is not rare that a course relies on different digital LMSs serving different purposes. In the use case proposed in this paper, an educator uses both Moodle and Graasp within the same course. Moodle is used as a repository for sharing course content, taking formative quizzes, and submitting assignments. Graasp is used for hands-on exercises with virtual laboratories and interactive applications. Adding a link on the Moodle course page pointing to the Graasp course space (and vice versa) easily enables centralized access to course resources. For a comprehensive overview of the overall course activity, the educator imports and exploits course-related activity traces from Moodle alongside those from Graasp. The aim is also to make those activity traces and outputs available to students. Therefore, they are accessible from within the course space, throughout the course. Furthermore, Graasp is also used as the publisher in this scenario, providing a sample repository of enablers that the educator can seamlessly add to the Graasp course space in order to import and visualize LA.

4.2 Contextual Integrator

The contextual integration of learning traces and related analytics constitutes an essential part of the approach brought forward in this paper. Instead of using a centralized LRS, learning traces are integrated into one or more of the course spaces. It is then up to other enablers (analyzers, presenters) to exploit those resources and expose the resulting output, also within the appropriate context.

Following the use case proposed in this paper, Graasp was chosen as a contextual integrator. Thus, activity traces imported from Moodle are added as usable resources—using a dedicated connector—to the corresponding course spaces on Graasp. Graasp does not impose a specific data format or data structures to resources added in a course space. Such specifications are delegated to the component or services exploiting the data. Furthermore, as shown in Figure 2, Graasp

can combine both interactive applications (e.g., a tool to write and execute Python code) alongside presenters that can be used to provide LA feedback to students.

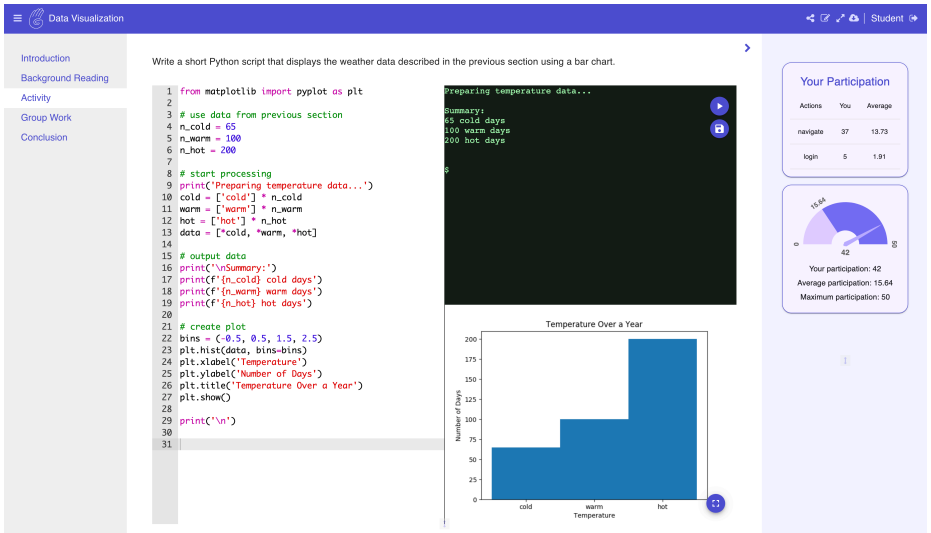

Fig. 2. A sample learning space on the Graasp platform. The main content—in this case, a coding exercise—is shown in the central panel. Learners can navigate between different activities using the menu on the left, while educators can provide LA feedback using presenters such as the *Your Participation* tool shown on the right column.

4.3 Moodle Connector

We built a connector component in order to import activity traces from Moodle into the course's dedicated space in Graasp. Moodle offers a rich Web service API that allows—after authentication with a Moodle user possessing the necessary privileges—the execution of various basic functions. The connector comprises three steps: (1) authenticate to Moodle, (2) select a relevant course and data of interest, and (3) import data to Graasp. Following these steps, we established a connection with our Moodle instance and invoked a Web service function. Through a custom Moodle plugin that adds a bespoke Web service function, we allow the export of course activity logs. Following a loosely coupled, standard-agnostic approach, the learning traces imported from Moodle are treated as any resource added to the course space in Graasp and are not expected to be treated or interpreted by Graasp in any special way. Figure 3 shows the user interface from the perspective of an educator using the connector in Graasp.

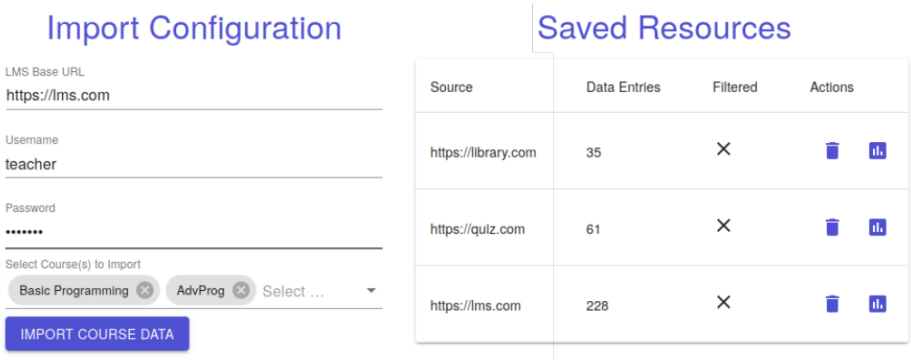

Fig. 3. The user interface of the connector to import Moodle data into Graasp. On the left, the configuration panel allows users to establish a connection with the targeted Moodle instance and select the course logs that are available to the user. On the right, a table summarizes the imported data available as resources for further processing by an analyzer or presenter.

4.4 Dashboard Presenter

Following the approach we present in this paper, the presenters are decoupled from the Moodle connector and allow the user to select the data to be displayed in order to maximize modular reuse. The choice of the appropriate dataset is delegated to the educator, who is provided with a description of the expected data source or attributes. The developed presenters are deployed as open-source plugins, which can be easily embedded into a Graasp learning space, as shown in Figure 2. Multiple presenters have been implemented to demonstrate the potential outputs of our approach. Figure 4 shows: (1) a stacked bar chart that can be used to show the distribution and the total number of activity traces alongside (2) a line chart with an interactive filter option, that also displays the total number of traces. These specific presenters combine two data sources: (1) the activity traces generated in Moodle, imported by the Moodle connector, and stored as a resource in the corresponding Graasp learning space, and (2) the activity traces generated by students while interacting with the Graasp learning space itself. Showing data coming from multiple sources side-by-side or combined in the same view helps educators attain a global overview of the activity that is happening within a course.

5 Conclusion and Future Work

In this paper, we propose an integrated, educator-driven, bottom-up approach to collect and exploit activity traces from different systems serving the same educational course. In the absence of a unique standard for LA and compared to traditional approaches focusing on particular use cases or combinations of data sources, our approach concentrates on enabling the discovery of meaningful LA

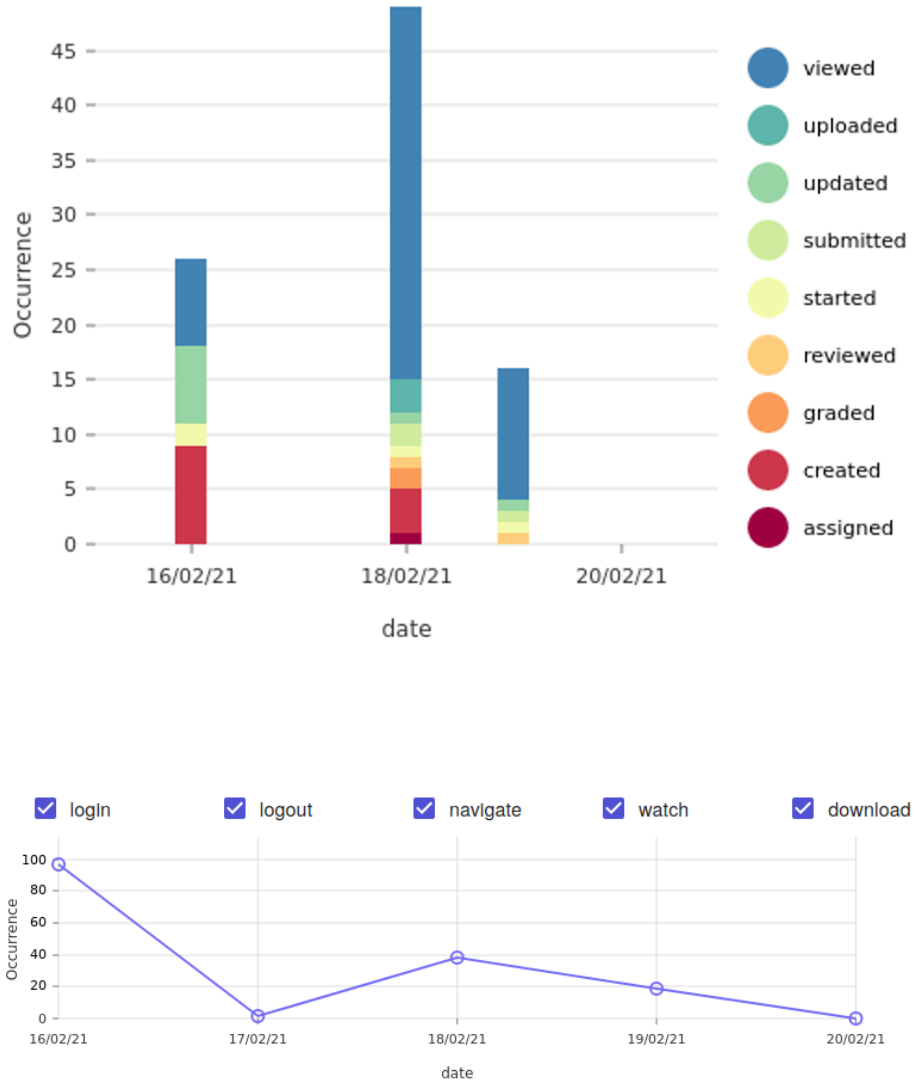

Fig. 4. The visualization result using two different presenters applied on distinct data sources (top: Moodle; bottom: Graasp) and showing how student activity is distributed over time. It is worth noting that the action verbs chosen serve the purpose of the example and it is up to the educator to define meaningful "verbs". The input data belongs to complementary LMSs used in the same course and serving different tasks. In this particular example, one can tell that the two LMSs involved were used mostly simultaneously but with varying intensity.

tools and the creation of a shared pool of reusable components. This is achieved by decoupling the import, semantic mapping, analysis, and visualization processes, leaving the choice and composition of components to the end-user.

We highlight the applicability of our approach with a proof-of-concept whereby a connector is developed to fetch learning activity traces from Moodle and embed them within Graasp for visualization using different presenters. By supporting cross-platform LA without the need for a standard or adaptations on the part of the learning platforms themselves, our goal is to support a bottom-up integration of LA that can more easily adapt to specific learning scenarios. The aim is to help educators attain a broader picture of student activity given that when multiple LMSs are used for the same course, aggregating the usage data from such complementary LMSs yields a more comprehensive activity overview, especially when viewed in context.

As a next step, we plan to develop an action mapper to allow educators to easily match action verbs coming from different data sources. Once in place, it will be possible to combine data sources with heterogeneous semantics in an analyzer, which brings our proof-of-concept closer to provide an overall assessment of student activity across learning tools applied in the presented use case. We will also evaluate the usability of our mapper and of a predefined set of mappings through participatory design workshops with educators. By involving educators in co-designing useful mappings, as well as analytics and visualization scenarios for presenters, we aim to seed a repository of enablers that can be published online and help kickstart a community around open, reusable, and adaptable LA solutions.

References

1. Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Mulder, R., Williams, D., Dawson, S., Lockyer, L.: A conceptual framework linking learning design with learning analytics. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. p. 329–338. LAK '16, Association for Computing Machinery, New York, NY, USA (2016), https://doi.org/10.1145/2883851.2883944
2. Clow, D.: An overview of learning analytics. Teaching in Higher Education 18(6), 683–695 (2013), https://doi.org/10.1080/13562517.2013.827653
3. Conde, M.Á., Hérnandez-García, Á., J. García-Peñalvo, F., Séin-Echaluce, M.L.: Exploring student interactions: Learning analytics tools for student tracking. In: Zaphiris, P., Ioannou, A. (eds.) Learning and Collaboration Technologies. pp. 50–61. Springer International Publishing, Cham (2015)
4. Flanagan, B., Ogata, H.: Learning analytics platform in higher education in japan. Knowledge Management & E-Learning: An International Journal 10(4), 469–484 (2018)
5. Gillet, D.: Personal learning environments as enablers for connectivist MOOCs. In: 2013 12th International Conference on Information Technology Based Higher Education and Training (ITHET). pp. 1–5. IEEE (2013), http://ieeexplore.ieee.org/document/6671026/

6. Greller, W., Drachsler, H.: Translating learning into numbers: A generic framework for learning analytics. Journal of Educational Technology & Society 15(3), 42–57 (2012)
7. Griffiths, C.: All about the lms. Learning Management Systems (2020)
8. Hernández-García, A., Conde, M.A., Chaparro-Peláez, J.: Learning analytics: A time to shine. In: Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality. p. 713–718. TEEM'20, Association for Computing Machinery, New York, NY, USA (2020), https://doi.org/10.1145/3434780.3436712
9. Hill, P.: Academic LMS Market Share: A view across four global regions (June 2017), https://eliterate.us/academic-lms-market-share-view-across-four-global-regions/
10. IMS Global Learning Consortium: Learning Measurement for Analytics Whitepaper. Tech. rep., IMS Global Learning Consortium, Inc., 801 International Parkway, 5th Floor, PMB #112, Lake Mary, FL 32746, USA (2013), https://www.imsglobal.org/sites/default/files/caliper/IMSLearningAnalyticsWP.pdf
11. Kitto, K., Cross, S., Waters, Z., Lupton, M.: Learning analytics beyond the lms: The connected learning analytics toolkit. In: Proceedings of the Fifth International Conference on Learning Analytics And Knowledge. p. 11–15. LAK '15, Association for Computing Machinery, New York, NY, USA (2015), https://doi.org/10.1145/2723576.2723627
12. Liu, D., Richards, D., Froissard, C., Atif, A.: Validating the effectiveness of the moodle engagement analytics plugin to predict student academic performance. In: 2015 Americas Conference on Information Systems, AMCIS 2015. pp. 1–10. Americas Conference on Information Systems (2015)
13. Murphy, M.P.A.: Covid-19 and emergency elearning: Consequences of the securitization of higher education for post-pandemic pedagogy. Contemporary Security Policy 41(3), 492–505 (2020), https://doi.org/10.1080/13523260.2020.1761749
14. Muslim, A., Chatti, M.A., Guesmi, M.: Open Learning Analytics: A Systematic Literature Review and Future Perspectives, pp. 3–29. Springer International Publishing, Cham (2020), https://doi.org/10.1007/978-3-030-41099-5_1
15. Rodríguez-Triana, M.J., Martínez-Monés, A., Villagrá-Sobrino, S.: Learning analytics in small-scale teacher-led innovations: Ethical and data privacy issues. Journal of Learning Analytics 3(1), 43–65 (Apr 2016), https://learning-analytics.info/index.php/JLA/article/view/4581
16. Romero, C., Ventura, S.: Educational data mining and learning analytics: An updated survey. WIREs Data Mining and Knowledge Discovery 10(3), e1355 (2020), https://onlinelibrary.wiley.com/doi/abs/10.1002/widm.1355
17. Rustici Software: xAPI Adopters. Rustici Software, LLC (March 2021), https://xapi.com/adopters/
18. Samuelsen, J., Chen, W., Wasson, B.: Integrating multiple data sources for learning analytics—review of literature. Research and Practice in Technology Enhanced Learning 14(1), 1–20 (2019)
19. Schwendimann, B.A., Rodríguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D., Dillenbourg, P.: Perceiving learning at a glance: A systematic literature review of learning dashboard research. IEEE Transactions on Learning Technologies 10(1), 30–41 (Jan 2017)
20. Shibani, A., Knight, S., Shum, S.B.: Contextualizable learning analytics design: A generic model and writing analytics evaluations. In: Proceedings of the 9th International Conference on Learning Analytics & Knowledge. p. 210–219. LAK19, Association for Computing Machinery, New York, NY, USA (2019), https://doi.org/10.1145/3303772.3303785

21. Siemens, G.: Learning and academic analytics. Learning and Knowledge Analytics 5 (2011)
22. Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S.B., Ferguson, R., Duval, E., Verbert, K., Baker, R., et al.: Open Learning Analytics: an integrated & modularized platform. Ph.D. thesis, Open University Press Doctoral dissertation (2011)
23. Strang, K.D.: Beyond engagement analytics: which online mixed-data factors predict student learning outcomes? Education and information technologies 22(3), 917–937 (2017)
24. Teasley, S.D.: Student facing dashboards: One size fits all? Technology, Knowledge and Learning 22(3), 377–384 (2017)
25. Viberg, O., Hatakka, M., Bälter, O., Mavroudi, A.: The current landscape of learning analytics in higher education. Computers in Human Behavior 89, 98 – 110 (2018), http://www.sciencedirect.com/science/article/pii/S0747563218303492
26. W3C: Activity streams 2.0. https://www.w3.org/TR/2017/REC-activitystreams-core-20170523/ (2017), accessed: 2021-21-06
27. Wilson, A., Watson, C., Thompson, T.L., Drew, V., Doyle, S.: Learning analytics: challenges and limitations. Teaching in Higher Education 22(8), 991–1007 (2017), https://doi.org/10.1080/13562517.2017.1332026
28. Wolff, A., Moore, J., Zdrahal, Z., Hlosta, M., Kuzilek, J.: Data literacy for learning analytics. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. p. 500–501. LAK '16, Association for Computing Machinery, New York, NY, USA (2016), https://doi.org/10.1145/2883851.2883864
29. Zilvinskis, J., Willis III, J., Borden, V.M.H.: An overview of learning analytics. New Directions for Higher Education 2017(179), 9–17 (2017), https://onlinelibrary.wiley.com/doi/abs/10.1002/he.20239