Designing Useful Learning Analytics: Developing an Adaptable LA Dashboard for the School of Informatics

Designing Useful Learning Analytics: Developing an Adaptable LA Dashboard for the School of Informatics

THE UNIVERSITY of EDINBURGH

Contributors:
Paul Boocock, Stefania I. Sandu, Xinyi Liu, Qiwen Liang, Heather Yorston, Aurora Constantin, Cristina Alexandru, Pavlos Andreadis.

Further Information:
For further information, please contact Pavlos at **pavlos.andreadis@ed.ac.uk**

Abstract:

Learning Analytics (LA) involves the collection and analysis of data for extracting insights and designing interventions that improve teaching and learning. The University of Edinburgh has laid out a LA policy that emphasises its strategic value. Despite the perceived benefits to our students, LA has not seen widespread adoption. We identify the ad-hoc development of customised solutions and the lack of availability of an adaptable infrastructure as key reasons. Our goal is the development of a LA Dashboard that works School-wide and adapts to the requirements of the individual user, computing infrastructure, and availability of data. This poster summarises the work done across a number of taught student projects in the School of Informatics, and lays down the direction of the ongoing work.

Why is Learning Analytics not Systematically Adopted in Higher Education?

As part of online and hybrid approaches to teaching, academic staff in our university have been using a variety of learning technology, spanning from virtual classroom systems (e.g. Blackboard Collaborate, Microsoft Teams, Zoom), virtual learning environments (e.g. Blackboard Learn, Moodle), discussion forums (e.g. Piazza), assessment and feedback systems (e.g. Learn assessments, Turnitin, CodeGrade, GradeScope, WebPA), blogging systems, electronic voting systems (e.g. TopHat). Some of these systems collect data on students and courses and offer statistics on a range of concerns like student engagement or attainment. Known as Learning Analytics, such an approach can greatly help evaluate approaches to teaching, course and programme design, and inform about improvements which could be made to enhance the students' learning and experience with university.

With this variability in sources of data [Juan Yang, et al. 2023], and given the diverse set of applications enabled by these learning technologies [Wong, Billy Tak-ming, and Kam Cheong Li 2020], Learning Analytics solutions have been primarily designed around specific [Viberg, Olga, et al. 2018], institution-wide, functionalities, rather than providing a tool that would allow teachers to investigate for developing their own learning solutions [Wong, Billy Tak-Ming et al. 2018][Joksimović, Srećko et al. 2019]. The sparse adoption of Learning Analytics by teachers, especially in higher education [Viberg, Olga, et al. 2018], might be related to this limitation.

Overview of Problem and Proposed Solution

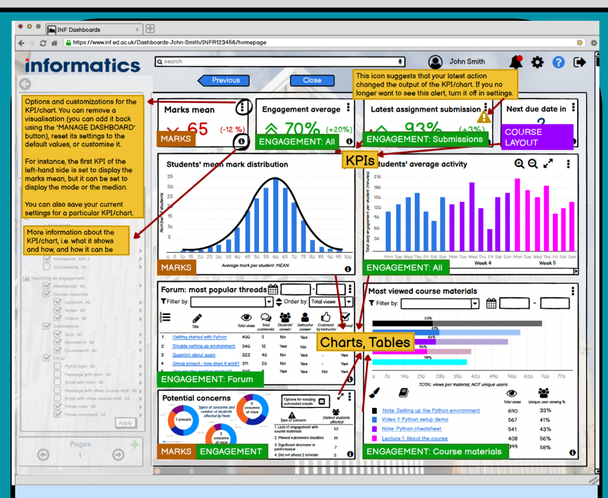
How can we address the problem with the adoption of Learning Analytics (LA) stemming from ad-hoc solutions not being applicable to changing data sources, and varied use cases?

> We build an adaptable system: capable of serving varied front-ends with no further work on the back-end; and capable of working off of evolving data sources with only the work required to ingest those into the existing data representations. This will allow it to adapt to new computing infrastructure within the School (especially changing data sources) and unique individual user requirements at the Front-End.

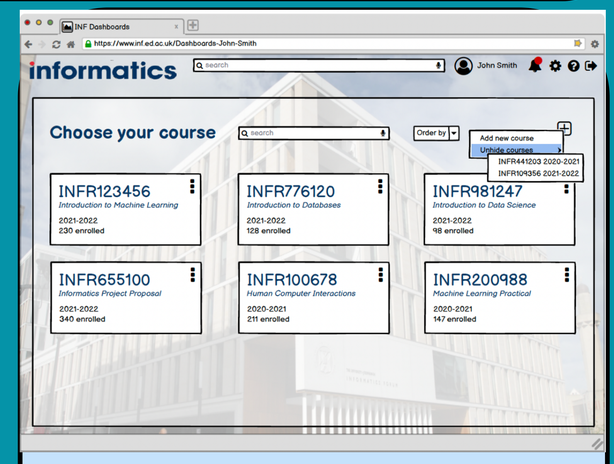
Accomplishing this requires that the system is built over data representations and abstractions of operations over the data with which one could design most, if not all, Learning Analytics dashboards.

A Data Analytics dashboard can be defined as "an interactive, historical, personalised, and analytical monitoring display that reflects students' learning patterns, status, performance, and interactions" [Park, Y. and Jo, I.H. 2015]. To that, we can add the ability to actively intervene on behalf of the student, in order to aid or prevent some negative outcome.

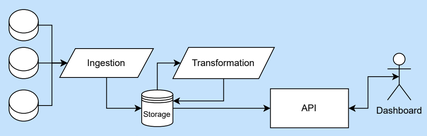
So it is with this mindset that we are looking at the available data sources in the School of Informatics, and the literature on Learning Analytics functionalities, with a focus on Informatics. While also producing some first designs of what dashboards such a system might serve. We have started work on designing the data representations and database, and are looking to continue in producing a general representation of operations that could be requested from a LA dashboard or other front end.



Surveying the School of Informatics, we produced a mock-up of a customisable interface, allowing for selectively including any of the functionalities requested by our colleagues.

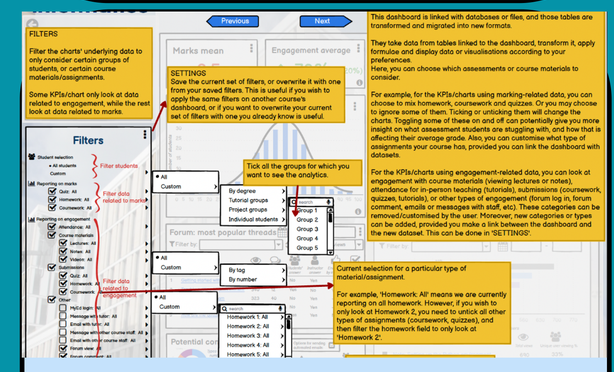


Users would be able to choose between pre-defined visualisations and toggle between various settings and functions, or they could create their own, using an API.



Abstracted view of a system supporting Learning Analytics Dashboards

- Building a data pipeline with flexible ingestion, to support multiple platforms, and an adaptable API for Learning Analytics Use Cases
- Need to offer ability to investigate data from multiple systems - find common entities and define them
- LA has a broad set of use cases, from many fields. We desire to offer an API that allows many people to inspect data - teachers, students, administrators
- Three layers - Ingestion, Transformation & Storage, API
- Ingest data using available APIs, Experience xAPI, Learn API, etc.
- Model data after ingestion, in a transformation layer, to conform to common entities, whilst storing platform specific information alongside
- Storage in suitable format - within a database or datalake [Medallion Design] (data lake offers broader query options and no vendor lock in)



Each implemented functionality would work over abstractions of data. Users could choose which subset of the source data feeds into the visualisations on multiple levels, such as by assignment (e.g. marks for all coursework), by learners (e.g., tutorial group 3), by topic, or by time frame.

Data Ingestion and Entity Mapping

Support for numerous data sources is a requirement, from the variety of learning technologies at use the University. Each system will contain data related to one of more core entities required for Learning Analytics. These systems are often unrelated and have no concept of each other, yet they contain data points which are important for Learning Analytics. Systems offering APIs, such as the Experience xAPI and Learn API, for data extraction but they vary in style and data they make available.

We propose to map the disparate data sets into common entities, which can then be joined together to form a dataset which users are able to power a multitude Learning Analytics Use Cases.

Learning Analytics Entities will include, but not be limited to: Students, Professors, Courses, Degrees, Assessments, Discussions, Comments.

Data Storage Formats and Technologies

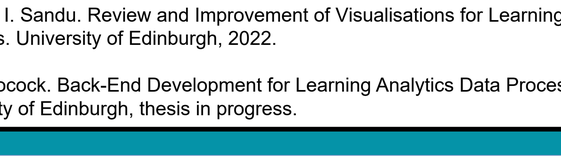
With data being ingested from numerous sources, in varying formats, it is sensible to store data in the raw state from the Data Source. This offers an unprocessed view of the data, allowing for validation and transformation.

This information can be stored in files, external to a database. It is increasingly common to leverage cloud object storage such as S3 [Amazon 2023] and store the data in an open standard format such as Apache Parquet [Databricks 2023].

To solve for a wide variety of questions and use cases, a data storage design optimised for analytics & business intelligence (BI) is required.

The medallion architecture classifies data into three stages of transformation, bronze, silver and gold [Microsoft 2023].

Each stage represents an increasingly optimised dataset however each stage is useful in isolation. Bronze offer trust in and reprocessing of the data. Silver is a clean but less aggregated dataset often useful for Data Science. Gold offer an optimised and aggregated dataset allowing for simple queries of modelled entities with established relationships.



Learning Analytics Interfaces (APIs)

The Gold, Analytics Ready, dataset can be loaded into a suitable database engine or programming library to be queried. A database engine will offer a SQL API over the data, allowing for a multitude of SQL based tools to be used, including Dashboard tools.

Additionally, by using open formats, programming libraries such as Pandas, Spark and Plotly can be used to manipulate and visualise the data.

User Requirements

To gather specifications for our system, we have a run a series of studies:

- A review of the LA Dashboards literature, with an emphasis on Informatics, identifying the common functionalities [Xinyi Liu 2023].
- First design and evaluation of a system built over the common functionalities identified in the literature [Qiwen Liang 2023].
- A review of the literature on types of visualisations and learning platforms, and subsequent interviewing of academic and supporting staff in the School of Informatics, followed by a thematic analysis of the findings, with the goal of evaluating these visualisations in terms of their usability and how they met the educational needs of our school. The screenshots in this poster are from the resulting dashboard prototype [Stefania I. Sandu 2022]

References:

Microsoft. What Is the Medallion Lakehouse Architecture? - Azure Databricks. 24 Feb. 2023, https://learn.microsoft.com/en-us/azure/databricks/lakehouse/medallion.

"Cloud Object Storage – Amazon S3 – Amazon Web Services." Amazon Web Services, Inc., https://aws.amazon.com/s3/. Accessed 30 May 2023.

"What Is Apache Parquet?" Databricks, 30 Oct. 2018, https://www.databricks.com/glossary/what-is-parquet.

A systematic meta-Review and analysis of learning analytics research Xu Du, Juan Yang ,Brett E.Shelton, Jui-Long Hung & Mingyan Zhang Pages 49-62 Received 16 May 2019, Accepted 14 Sep 2019, Published online: 23 Sep 2019. https://www.tandfonline.com/doi/full/10.1080/0144929X.2019.1669712

Wong, Billy Tak-ming, and Kam Cheong Li. "A review of learning analytics intervention in higher education (2011–2018)." Journal of Computers in Education 7.1 (2020): 7-28. https://link.springer.com/article/10.1007/s40692-019-00143-7

Wong, Billy Tak-Ming, Kam Cheong Li, and Samuel Ping-Man Choi. "Trends in learning analytics practices: a review of higher education institutions." Interactive Technology and Smart Education (2018). https://www.emerald.com/insight/content/doi/10.1108/ITSE-12-2017-0065/full/html

Joksimović, Srećko, Vitomir Kovanović, and Shane Dawson. "The journey of learning analytics." HERDSA Review of Higher Education 6 (2019): 27-63. https://www.researchgate.net/profile/Vitomir-Kovanovic/publication/335678985_The_Journey_of_Learning_Analytics/links/5d73a05c299bf1cb808ebcae/The-Journey-of-Learning-Analytics.pdf

Viberg, Olga, et al. "The current landscape of learning analytics in higher education." Computers in Human Behavior 89 (2018): 98-110. https://www.sciencedirect.com/science/article/pii/S0747563218303492

Park, Y. and Jo, I.H. (2015). Development of the learning analytics dashboard to support students' learning performance. Journal of Universal Computer Science, 21(1), p.110.

Learning analytics. the University of Edinburgh. https://www.ed.ac.uk/academic-services/policies-regulations/learning-and-assessment/learning-analytics Accessed 2023-05-10

Xinyi Liu. Data Sourcing for Learning Analytics. University of Edinburgh, 2023.

Qiwen Liang. Data Pipeline for Learning Analytics. University of Edinburgh, 2023.

Stefania I. Sandu. Review and Improvement of Visualisations for Learning Analytics. University of Edinburgh, 2022.

Paul Boocock. Back-End Development for Learning Analytics Data Processing. University of Edinburgh, thesis in progress.