

System-wide data analytics for learning analytics and data-driven decision making in learning software

Jan Renz¹, Max Klenk² and Christoph Meinel³

Abstract: We perform learning analytics to provide all stakeholder in the e-learning environment with insights, which are relevant for them. To archive this, usually a fix set of metrics is shown in a dashboard. Direct access to the source data is normally not possible. This work examines the technical possibilities to allow analyses on the relevant data in distributed learning applications. In the context of distributed systems, like microservice-based architectures, the data is stored in decentral silos, which requires to merge them before the analysis can start. We additionally research how this approach can be combined with a dedicated learning analytics store and which tools would help to archive this.

Keywords: Microservices, Open-Source, Learning-Analytics, MOOCs

¹ Hasso Plattner Institut, Universität Potsdam, jan.renz@hpi.de

² Hasso Plattner Institut, Universität Potsdam, max.klenk@student.hpi.de

³ Hasso Plattner Institut, Universität Potsdam, christoph.meinel@hpi.de

Systemweite Datenanalyse für Learning Analytics und datengetriebene Entscheidungen in Lernumgebungen

Jan Renz⁴, Max Klenk⁵ und Christoph Meinel⁶

Abstract: Wir betreiben Learning Analytics, um alle Akteure und Stakeholder im E-Learning Umfeld mit den für sie relevanten Daten zu versorgen. Im Kontext von verteilten Systemen wie mit Microservices gebauten Lernarchitekturen liegen die Anwendungsdaten oft dezentral in Silos, sodass diese für die Analyse erst zusammengebracht werden müssen. Der direkte Zugriff auf die Quell-Daten ist oft nicht möglich. Diese Arbeit untersucht die technischen Möglichkeiten, relevante Daten aus verteilten Lernanwendungen über einen zentralen Zugang zugreif- und auswertbar zu machen.

Keywords: Microservices, Open-Source, Learning-Analytics, MOOCs

1 Einleitung

In einer Microservice-basierten Architektur gibt es keinen zentralen Ort, an dem alle anfallenden Daten gespeichert werden, da alle Dienste sich um das Persistieren ihrer eignen, Domänen-spezifischen Daten kümmern. Für die Implementierung einer Learning Analytics Umgebung entstehen dadurch zwei Herausforderungen: Einen geeigneten Ansatz für das Sammeln und Extrahieren der Daten finden, sowie die Umsetzung einer Architektur, die das Speichern, Auswerten und Abfragen der Daten ermöglicht. Mit diesen Daten können dann den verschiedenen Stakeholdern Anwendungen und Oberflächen zur Verfügung gestellt werden. Da die benötigten Daten entweder über mehrere Dienste verteilt sind oder bisher noch in keinem Dienst erfasst wurden (insbesondere Nutzer-Interaktionsdaten), verfügt die HPI MOOC-Plattform über einen dedizierten Learning Analytics Dienst [Re16]. Wie in [Ro19a] erläutert, funktioniert dieses Konzept gut und ist skalierbar, stößt aber auch an Grenzen: (a) In verteilten Systemen ist es aufwendig, die Analytics Daten mit weiteren Daten zu korrelieren die nicht bei der Erfassung berücksichtigt wurden. (b) Erweiterungen der Dashboards sind zeit- und ressourcenaufwendig, da nur die Entwickler darauf Zugriff haben. (c) Exporte bieten nur eine Momentaufnahme der Daten; diese regelmäßig zu updaten und in lokale Analysen einfließen zu lassen ist aufwendig. (d) Um Exporte lokal zu analysieren, muss jeder eigenständig eine neue Arbeitsumgebung schaffen und Analyse-Tools installieren. (e) Die lokalen Analysen sind schwer mit anderen zu teilen oder gar in das Live-System zu übernehmen.

⁴ Hasso Plattner Institut, Universität Potsdam, jan.renz@hpi.de

⁵ Hasso Plattner Institut, Universität Potsdam, max.klenk@student.hpi.de

⁶ Hasso Plattner Institut, Universität Potsdam, christoph.meinel@hpi.de

2 Erweiterung der Architektur

Um diese Probleme zu lösen und eine interaktivere Nutzung der Daten zu befördern wird die bestehende HPI MOOC-Plattform so erweitert, dass die verschiedenen Datenquellen vereinheitlicht werden und eine Online-Arbeitsumgebung geschaffen wird, die es erlaubt, Analysen direkt auf diesen Daten zu entwickeln.

Um die Analytics-Daten mit weiteren Daten zu kombinieren, die nicht schon in dem Learning Analytics Data Warehouse berücksichtigt sind, wird ein Zugriff auf die Datenbanken der verschiedenen Microservices und ggf. auch auf externe Tools [Ro19b] benötigt. Tools zur **Query Federation** wie Apache Presto erlauben es, in einer Anfrage Datenquellen zu nutzen. Um den Export der Daten und die lokale Analyse weitgehend zu vermeiden, wird zusätzlich eine gehostete **kollaborative Arbeitsumgebung** via Jupyter, Metabase oder Zeppelin geschaffen.

Möchte man bspw. Informationen über das Alter der Kursteilnehmer auswerten, so liegen die Daten über die Kurszugehörigkeit im Kursservice, die des Alters hingegen im Benutzerservice. Die Performance einer solchen Auswertung ist dabei mit Query Federation deutlich höher als beim Abfragen der gleichen Daten aus den Services per REST. Ein **Metrikproxy** wird als zusätzliche Abstraktionsschicht und zum Caching eingesetzt.

Diese drei Erweiterungen adressieren die identifizierten Unzulänglichkeiten des Learning Analytics Dienstes der HPI MOOC-Plattform, bringen aber auch eine eigene Komplexität mit sich und erhöhen Wartungsaufwände. Die technische Verfügbarkeit und Machbarkeit lassen auch noch keine Rückschlüsse auf die Nutzung und Akzeptanz zu. Daher muss beobachtet werden, wie die neuen Möglichkeiten genutzt werden. Dazu gehört neben einem Performancemonitoring und der Evaluation der Nutzung auch die Beantwortung der Frage, inwieweit die potentiellen Nutzer an die Hand genommen werden müssen und welche Maßnahmen dazu führen, dass ein zunehmend kollaborativer Arbeitsmodus eingenommen wird.

Literaturverzeichnis

- [Re16] Renz, J.; Navarro-Suarez, G.; Sathi, R.; Staubitz, T.; Meinel, C.: Enabling schema agnostic learning analytics in a service-oriented MOOC platform. In: Proceedings of the Third (2016) ACM Conference on Learning@ Scale.
- [Ro19a] Rohloff, T.; Renz, J.; Suarez, G. N.; Meinel, C.: A Ubiquitous Learning Analytics Architecture for a Service-Oriented MOOC Platform. In (Calise M.; Delgado Kloos, C.; Reich, J.; Ruiperez-Valiente, J. A.; Wirsing, M., Hrsg.): Digital Education: At the MOOC Crossroads Where the Interests of Academia and Business Converge. Springer International Publishing, Cham, S. 162–171, 2019, ISBN: 978-3-030-19875-6.
- [Ro19b] Rohloff, T.; Oldag, S.; Renz, J.; Meinel, C.: Utilizing web analytics in the context of learning analytics for large-scale online learning. In: 2019 IEEE Global Engineering Education Conference (EDUCON). IEEE, S. 296–305, 2019.