DISSERTATION

Learning Analytics at Scale

Supporting Learning and Teaching in MOOCs with Data-Driven Insights

by

Tobias Rohloff

Hasso Plattner Institute for Digital Engineering
at the University of Potsdam

Supervisor:
Prof. Dr. Christoph Meinel (Hasso Plattner Institute, University of Potsdam)

Reviewers:
Prof. Dr. Hendrik Drachsler (Goethe University Frankfurt am Main and DIPF)
Prof. Dr. Marco Kalz (Heidelberg University of Education)

Date of Submission: May 20, 2021
Date of Disputation: October 15, 2021

Published online on the
Publication Server of the University of Potsdam:
https://doi.org/10.25932/publishup-52623
https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-526235

Tobias Rohloff. *Learning Analytics at Scale: Supporting Learning and Teaching in MOOCs with Data-Driven Insights*. Dissertation submitted in partial fulfillment of the requirements for the academic degree of "Doctor rerum naturalium" (Dr. rer. nat.) in the scientific discipline "Internet Technologies and Systems" to the Digital Engineering Faculty at the Hasso Plattner Institute, University of Potsdam. May 20, 2021.

Abstract

Digital technologies are paving the way for innovative educational approaches. The learning format of Massive Open Online Courses (MOOCs) provides a highly accessible path to lifelong learning while being more affordable and flexible than face-to-face courses. Thereby, thousands of learners can enroll in courses mostly without admission restrictions, but this also raises challenges. Individual supervision by teachers is barely feasible, and learning persistence and success depend on students' self-regulatory skills. Here, technology provides the means for support. The use of data for decision-making is already transforming many fields, whereas in education, it is still a young research discipline. Learning Analytics (LA) is defined as the measurement, collection, analysis, and reporting of data about learners and their learning contexts with the purpose of understanding and improving learning and learning environments. The vast amount of data that MOOCs produce on the learning behavior and success of thousands of students provides the opportunity to study human learning and develop approaches addressing the demands of learners and teachers.

The overall purpose of this dissertation is to investigate the implementation of LA at the scale of MOOCs and to explore how data-driven technology can support learning and teaching in this context. To this end, several research prototypes have been iteratively developed for the HPI MOOC Platform. Hence, they were tested and evaluated in an authentic real-world learning environment. Most of the results can be applied on a conceptual level to other MOOC platforms as well. The research contribution of this thesis thus provides practical insights beyond what is theoretically possible. In total, four system components were developed and extended:

(1) The Learning Analytics Architecture: A technical infrastructure to collect, process, and analyze event-driven learning data based on schema-agnostic pipelining in a service-oriented MOOC platform. (2) The Learning Analytics Dashboard for Learners: A tool for data-driven support of self-regulated learning, in particular to enable learners to evaluate and plan their learning activities, progress, and success by themselves. (3) Personalized Learning Objectives: A set of features to better connect learners' success to their personal intentions based on selected learning objectives to offer guidance and align the provided data-driven insights about their learning progress. (4) The Learning Analytics Dashboard for Teachers: A tool supporting teachers with data-driven insights to enable the monitoring of their courses with thousands of learners, identify potential issues, and take informed action.

For all aspects examined in this dissertation, related research is presented, development processes and implementation concepts are explained, and evaluations are conducted in case studies. Among other findings, the usage of the learner dashboard in combination with personalized learning objectives demonstrated improved certification rates of 11.62% to 12.63%. Furthermore, it was observed that the teacher dashboard is a key tool and an integral part for teaching in MOOCs. In addition to the results and contributions, general limitations of the work are discussed—which altogether provide a solid foundation for practical implications and future research.

Zusammenfassung

Digitale Technologien sind Wegbereiter für innovative Bildungsansätze. Das Lernformat der Massive Open Online Courses (MOOCs) bietet einen einfachen und globalen Zugang zu lebenslangem Lernen und ist oft kostengünstiger und flexibler als klassische Präsenzlehre. Dabei können sich Tausende von Lernenden meist ohne Zulassungsbeschränkung in Kurse einschreiben, wodurch jedoch auch Herausforderungen entstehen. Eine individuelle Betreuung durch Lehrende ist kaum möglich und das Durchhaltevermögen und der Lernerfolg hängen von selbstregulatorischen Fähigkeiten der Lernenden ab. Hier bietet Technologie die Möglichkeit zur Unterstützung. Die Nutzung von Daten zur Entscheidungsfindung transformiert bereits viele Bereiche, aber im Bildungswesen ist dies noch eine junge Forschungsdisziplin. Als Learning Analytics (LA) wird das Messen, Erfassen, Analysieren und Auswerten von Daten über Lernende und ihren Lernkontext verstanden, mit dem Ziel, das Lernen und die Lernumgebungen zu verstehen und zu verbessern. Die riesige Menge an Daten, die MOOCs über das Lernverhalten und den Lernerfolg produzieren, bietet die Möglichkeit, das menschliche Lernen zu studieren und Ansätze zu entwickeln, die den Anforderungen von Lernenden und Lehrenden gerecht werden.

Der Schwerpunkt dieser Dissertation liegt auf der Implementierung von LA für die Größenordnung von MOOCs und erforscht dabei, wie datengetriebene Technologie das Lernen und Lehren in diesem Kontext unterstützen kann. Zu diesem Zweck wurden mehrere Forschungsprototypen iterativ für die HPI-MOOC-Plattform entwickelt. Daher wurden diese in einer authentischen und realen Lernumgebung getestet und evaluiert. Die meisten Ergebnisse lassen sich auf konzeptioneller Ebene auch auf andere MOOC-Plattformen übertragen, wodurch der Forschungsbeitrag dieser Arbeit praktische Erkenntnisse über das theoretisch Mögliche hinaus liefert. Insgesamt wurden vier Systemkomponenten entwickelt und erweitert:

(1) Die LA-Architektur: Eine technische Infrastruktur zum Sammeln, Verarbeiten und Analysieren von ereignisgesteuerten Lerndaten basierend auf einem schemaagnostischem Pipelining in einer serviceorientierten MOOC-Plattform. (2) Das LA-Dashboard für Lernende: Ein Werkzeug zur datengesteuerten Unterstützung der Selbstregulierung, insbesondere um Lernende in die Lage zu versetzen, ihre Lernaktivitäten, ihren Fortschritt und ihren Lernerfolg selbst zu evaluieren und zu planen. (3) Personalisierte Lernziele: Eine Reihe von Funktionen, um den Lernerfolg besser mit persönlichen Absichten zu verknüpfen, die auf ausgewählten Lernzielen basieren, um Leitlinien anzubieten und die bereitgestellten datengetriebenen Einblicke über den Lernfortschritt darauf abzustimmen. (4) Das LA-Dashboard für Lehrende: Ein Hilfsmittel, das Lehrkräfte mit datengetriebenen Erkenntnissen unterstützt, um ihre Kurse mit Tausenden von Lernenden zu überblicken, mögliche Probleme zu erkennen und fundierte Maßnahmen zu ergreifen.

Für alle untersuchten Aspekte dieser Dissertation werden verwandte Forschungsarbeiten vorgestellt, Entwicklungsprozesse und Implementierungskonzepte erläutert und Evaluierungen in Fallstudien durchgeführt. Unter anderem konnte durch den Einsatz des Dashboards für Lernende in Kombination mit personalisierten Lernzielen verbesserte Zertifizierungsraten von 11,62% bis 12,63% nachgewiesen werden. Außerdem wurde beobachtet, dass das Dashboard für Lehrende ein entscheidendes Werkzeug und ein integraler Bestandteil für die Lehre in MOOCs ist. Neben den Ergebnissen und Beiträgen werden generelle Einschränkungen der Arbeit diskutiert, die insgesamt eine fundierte Grundlage für praktische Implikationen und zukünftige Forschungsvorhaben schaffen.

Für Opa, in liebevoller Erinnerung.

Acknowledgements

I would like to express my gratitude to all those who have made this work possible and supported me along the way.

First of all, I would like to thank my supervisor Prof. Dr. Christoph Meinel, for his guidance, ideas, and the opportunity to pursue a doctoral degree in a socially valuable project.

I would also like to thank the whole openHPI team and my fellow researchers, in particular Jan Renz, Christian Willems, Tom Staubitz, Max Bothe, Christiane Hagedorn, Franz Liedke, Jan Graichen, Malte Swart, Ralf Teusner, Sebastian Serth, Matthias Wiesner, Gerardo Navarro Suarez, and Karen von Schmieden. Thanks for your support, feedback, co-authorship, and technical expertise.

A very special thanks goes to my master's students Dominic Sauer (in the meantime colleague) and Sören Oldag, with whom I have developed, examined, and published parts of this thesis.

Also, I would like to thank the openSAP team, especially Florian Schwerer, Clemens Link, and Nathanael Schenk, for your feedback and the opportunity to deploy and study my research prototypes on your platform.

Another thanks goes to my student assistants Ben-Noah Engelhaupt, Bjarne Sievers, Sandro Speh, and Jonas Müller for their technical contributions.

Last but not least, I would like to thank my loved ones who have supported me in all circumstances of life, especially my parents, grandparents, and Sophie.

Thank you all.

List of Pre-Publications

The following pre-published papers have contributed to the completion of this thesis and substantial parts have been reused.

1. T. Rohloff, M. Bothe, J. Renz, and C. Meinel. "Towards a Better Understanding of Mobile Learning in MOOCs". In: *Proceedings of 2018 IEEE Learning With MOOCS*. LWMOOCS '18. IEEE, 2018, pp. 1–4. DOI: 10.1109/LWMOOCS.2018.8534685
2. T. Rohloff and C. Meinel. "Towards Personalized Learning Objectives in MOOCs". In: *Lifelong Technology-Enhanced Learning*. EC-TEL '18. Springer International Publishing, 2018, pp. 202–215. DOI: 10.1007/978-3-319-98572-5_16
3. T. Rohloff, J. Renz, G. N. Suarez, and C. Meinel. "A Ubiquitous Learning Analytics Architecture for a Service-Oriented MOOC Platform". In: *Digital Education: At the MOOC Crossroads Where the Interests of Academia and Business Converge*. EMOOCs '19. Springer International Publishing, 2019, pp. 162–171. DOI: 10.1007/978-3-030-19875-6_19
4. T. Rohloff, S. Oldag, J. Renz, and C. Meinel. "Utilizing Web Analytics in the Context of Learning Analytics for Large-Scale Online Learning". In: *Proceedings of the 2019 IEEE Global Engineering Education Conference*. EDUCON '19. IEEE, 2019, pp. 296–305. DOI: 10.1109/EDUCON.2019.8725118
5. T. Rohloff, D. Sauer, and C. Meinel. "On the Acceptance and Usefulness of Personalized Learning Objectives in MOOCs". In: *Proceedings of the Sixth ACM Conference on Learning @ Scale*. L@S '19. Association for Computing Machinery, 2019, 4:1–4:10. DOI: 10.1145/3330430.3333624
6. T. Rohloff, D. Sauer, and C. Meinel. "Student Perception of a Learner Dashboard in MOOCs to Encourage Self-Regulated Learning". In: *Proceedings of the 2019 IEEE International Conference on Engineering, Technology and Education*. TALE '19. IEEE, 2019, pp. 1–8. DOI: 10.1109/TALE48000.2019.9225939
7. T. Rohloff, D. Sauer, and C. Meinel. "Students' Achievement of Personalized Learning Objectives in MOOCs". In: *Proceedings of the Seventh ACM Conference on Learning @ Scale*. L@S '20. Association for Computing Machinery, 2020, pp. 147–156. DOI: 10.1145/3386527.3405918
8. T. Rohloff, K. von Schmieden, and C. Meinel. "Students' Satisfaction of a Design Thinking MOOC with Personalized Learning Objectives". In: *Proceedings of 2020 IEEE Learning With MOOCs*. LWMOOCS '20. IEEE, 2020, pp. 37–41. DOI: 10.1109/LWMOOCS50143.2020.9234349

Contents

1 General Introduction

"Learning analytics has potential to dramatically impact the existing models of education and to generate new insights into what works and what does not work in teaching and learning."

— George Siemens [139]

Learning means development and thus learning tools have to develop as well. To get one step closer to Siemens's vision, this doctoral thesis investigates the implementation of Learning Analytics (LA) at the scale of Massive Open Online Courses (MOOCs) and explores how data-driven technology can support learning and teaching in this context. For this purpose, several research prototypes are developed for the HPI MOOC Platform, which enables to test and evaluate them in an authentic learning environment. As an introduction to the overall topic, this chapter explains the thesis' background and research motivation and presents its main research aims and questions. It is followed by a summary of key definitions and terms used in this thesis. Afterward, the research context is introduced. The chapter concludes by outlining the structure of the thesis.

1.1 Research Motivation and Background

The intersection of education and technology has accompanied me throughout my life, as the son of a teacher and an engineer. During my graduate studies, I joined the openHPI project to work on the technical side of MOOCs in 2014. In the same year the first "ACM Conference on Learning at Scale" took place, which emerged from the momentum of MOOCs, but also with the intention to change and improve learning and teaching beyond this format when done at scale [40]. Originally set out with the utopian goal of revolutionizing higher education, MOOCs have since found their reasonable applications in the global education landscape [103]. They provide a highly accessible way to lifelong learning while being more affordable and flexible than in-person courses. These benefits are achieved through technological innovations, but they have a tremendous impact on the way teaching and learning are performed, and thus raise challenges. Teachers oversee a learning community of typically tens of thousands of students with diverse motivations and backgrounds, spread across the globe. Students need to self-regulate their learning, as this seems to be essential for persistence and success [3, 71]. And in the course of this, communication, guidance, and feedback between all stakeholders usually only take place asynchronously. Here, too, technology provides the means for support.

"Information is the oil of the 21st century, and analytics is the combustion engine," stated Peter Sondergaard in a speech in 2011, who was senior vice-president of Gartner at this time. This popular metaphorical phrase has been much discussed, with all the concerns [155] and opportunities [7] lying hidden in data. The use of data and evidence for decision-making is already transforming many fields, whereas in education, it is still a young research discipline [144]—the main assembly in form of the “International Conference on Learning Analytics & Knowledge” celebrated its 10^{th} anniversary last year [73]. Learning analytics has the potential to better understand and optimize the learning and teaching in the digital realm for all stakeholders [140]—by revealing insights into the past, analyzing why things did happen, and advising on possible outcomes [145]—when privacy and ethical issues are addressed [30]. The results can be transformative for the entire education system [139], when adopted at scale [27, 38].

The vast amount of data that MOOCs produce on the behavior and success of thousands of students provides the opportunity to study human learning and develop approaches addressing the demands of learners and instructors. Hence, MOOCs are “gold mines” for learning analytics [54], but most MOOCs ignore this potential to facilitate awareness, self-regulation, and personalization with data-driven methods [32]. The first groundwork on learning analytics in the openHPI project has started in 2015 [105] leading to the initiation of this thesis in 2016. At that time, there were only a few published case studies that applied LA to MOOCs, as well as conceptual works. Even to this day, this has only improved slightly. The setup of having both the platform’s development and the educational content’s production and facilitation implemented and managed by the openHPI team has proven to be a unique opportunity to mine the inherent data treasure and thus create added value for all stakeholders. Therefore, the motivation for this thesis is to enable general LA capabilities for the HPI MOOC Platform and implement use cases for platform owners, course teachers, learners, and researchers, while contributing to the body of knowledge on the implementation of LA in MOOCs.

1.2 Research Aims and Main Research Questions

The overall purpose of this work is to explore technical concepts for the integration and application of learning analytics in MOOCs to support learners and teachers with data-driven insights. To this end, answers to the following research questions are presented throughout this thesis:

Research Question 1: How can learning analytics be enabled at the scale of MOOCs?

Research Question 2: How can data-driven insights support learning in MOOCs?

Research Question 3: How can data-driven insights support teaching in MOOCs?

In order to examine these topics from the perspective of various more detailed aspects, further sub-questions are derived and investigated in the subsequent chapters. In the process, several prototypes are iteratively developed and tested in case studies. Thereby, for all aspects examined in this thesis, related research is presented, development processes and implementation concepts are explained, and evaluations are conducted based on the derived research questions with mixed-methods, e. g., platform data analyses, questionnaires, and A/B/n tests. This is performed entirely within the context of the HPI MOOC Platform, which allows us to base all results on real-world data and authentic learning experiences. Most of the results can be applied on a conceptual level to other MOOC platforms as well. The research contribution of this thesis thus provides practical insights beyond what is theoretically possible. This also includes technical requirements and limitations, as well as challenges in development processes in the everyday operation of a MOOC platform. This can help other platform providers and researchers to reproduce, adapt, and further enhance our approaches and findings.

1.3 Key Terms and Definitions

This section provides an overview about the main terms and definitions that are used in this thesis. First, MOOCs are explained with their history and characteristics (Subsection 1.3.1). Second, learning analytics is introduced in Subsection 1.3.2. Further terms and concepts, which are only relevant in certain chapters of this thesis, are explained there accordingly.

1.3.1 Massive Open Online Courses

In the history of e-learning, the MOOC phenomenon has its roots in the virtual learning environment (VLE) and the open course ware (OCW) movement. VLEs are web-based platforms that originated in the mid 90s, which support the digital aspects of a class by providing information, learning material, quizzes, etc. about a course [167]. In the late 90s, the Massachusetts Institute of Technology (MIT) and other universities began to provide recordings and material of some of their courses online, with free access for everyone. This was the start of the OCW movement, with its mission of opening up higher education to the general public. In the same year, the UNESCO held a forum about the impact of OCW for higher education in developing countries, from which the term open educational resources (OER) derived for digital learning materials that can be used freely for teaching [22]. These ideas were bundled by MOOCs. The term was created by Dave Cormier and Bryan Alexander in Canada for the open online course “Connectivism & Connective Knowledge Course 2008”, which was designed by George Siemens and Stephen Downes at the University of Manitoba [78]. Siemens [141] describes MOOCs as “a middle ground for teaching and learning between the highly organized and structured classroom environment and the chaotic open web of fragmented information.” For a better understanding it is useful to dissolve the acronym:

Massive: MOOCs reach enrollments with thousands of learners, with some exceeded more than 100,000 registrants. This enables participants to form sub-networks around their native language or geographical location and results in a highly diverse learning community enriching the learning process through multiple perspectives and ideas.

Open: Learners can access the course content and participate for free without fees, as long as an Internet connection is provided.

Online: The course is accessed exclusively online, which is enabled by web-based platforms and modern web technologies, as well as the support of mobile devices.

Course: MOOCs have an official start and end date, hence the social interaction by the learning community is focused. Also, they are structured and sequenced to ensure that all learners are on the same page. Still, learners are free to decide when and where they want to browse the resources. After the end of a MOOC, its content is mostly still accessible in an archived mode.

Additionally, MOOCs are commonly classified as Extension MOOCs (xMOOCs) and Connectivism MOOCs (cMOOCs) [141], although for simplification, in this thesis the term MOOC always refers to the xMOOC concept unless explicitly stated otherwise:

xMOOCs: These MOOCs are derived from traditional university courses, with typical elements like video lectures, quizzes, and forums. It follows the structure and content given by the teachers. Student interactions are limited to the learning process.

cMOOCs: The structure of these MOOCs is more loosely. Based on a main topic and a time plan the teachers give some resources to study. Then, the participants can self-reliant submit more material that can be discussed. This is intended to connect the student and knowledge in networks, to offer a more personalized learning experience.

A major breakthrough for MOOCs was the artificial intelligence course by Sebastian Thrun from Stanford University in 2011 with an astonishing number of over 150,000 learners from around the globe. The peak of the MOOC hype was reached in 2012, where the New York Times declaring it as "the year of the MOOC" [96]. In the following, the label MOOC was attached to just about anything that was some kind of online course, which gradually blurred the concept. In particular, the push by commercial providers has also pulled more and more content behind a paywall, calling into question the openness of many courses. After the hype flattened out, MOOCs were declared dead several times, but the current figures speak a different language [136]. The form and use of MOOCs has perhaps become more multifaceted and shifted towards lifelong learning [64, 103], but they remain a global phenomenon.

1.3.2 Learning Analytics

Modern technology paving the way for new tools in education, like virtual learning environments [12, 28] and ubiquitous computing devices [83, 165]. Nevertheless, the most influential technological aspect that can improve the learners progress and outcome is something they do not directly interact with: data analysis, which leads to the term Learning Analytics (LA). Data-driven decision-making can improve organizational output and productivity, which is already a common practice in many disciples like healthcare, government, or business [144]. In the domain of gaining insights from large data sets of learners while they are interacting with educational software and online learning, two distinct research communities were established: learning analytics and educational data mining (EDM) [142]. Thereby, EDM puts its focus on a technological perspective (data-driven analytics) and LA on a pedagogical perspective (learner-focused analytics), in order to monitor and improve teaching and learning [17, 37].

Learning analytics was defined from many different perspectives since it has a relationship to a variety of disciplines, including (1) learning and educational sciences, which compile pedagogical foundations and rationales, as well as theories and models about teaching and learning; (2) data sciences, which provide mathematical and statistical methods to process, analyze, and visualize learning data; and (3) computer sciences, which provide human-centered and innovative means to design and develop supportive learning tools [122, 145]. The most popular and accepted definition was formulated in the context of the first International Conference on Learning Analytics & Knowledge (LAK 2011) as "learning analytics is the measurement, collection, analysis, and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" [140]. From the perspective of the research field of the technology-enhanced learning (TEL), LA serves to develop methods for analyzing data collected from educational settings in order to support the learning experience [17].

Learning analytics is of particular interest for the study of MOOCs, and simultaneously the tremendous quantities of data that MOOCs produce serve to further advance the research field [32, 54]. Thereby, learning analytics research studies can be organized into three levels of impact according to Shum [138]: (1) micro-level analytics, to capture and analyze fine-grained process data of individual learners or learner groups, which in turn are then utilized for the learners themselves and their success, e. g., through personalized feedback or encouraged self-reflection; (2) meso-level analytics, to support decision-making and optimize processes with integrated data at the institutional level for educational organizations or platforms; and (3) macro-level analytics, to enable cross-institutional insights with combined data sources to inform and transform the community, organizations, and regulatory authorities.

Applied to MOOCs, learning analytics enables a variety of opportunities to support the diverse interests and needs of all stakeholders [31, 47], for which a few examples are given as follows. Next to analytics about completion rates and dropout predictions [21, 94], a focus on learners (micro-level) can support personalization, feedback, assessment, recommendation, awareness, and self-reflection leading to more effective MOOCs [17], e. g., by means of tools such as personalized learner dashboards [23, 57, 98]. Also, LA can provide insights for platform providers (meso-level) to improve the learning environment, e. g., through a better understanding of the learners' population [15] and demographics [59], or learning behavior [119]. Last, studies with combined data from different platforms (macro-level) can inform the research and practitioner communities, e. g., by analyzing trends across global and regional providers with different cultural backgrounds [123] or exploring interventions for closing global achievement gaps [66].

1.4 Research Context

After the initial success of different MOOC pilot projects, it did not last long until the first platforms and long-time offers emerged. The MIT and Harvard University combined their efforts and founded the not-for-profit edX platform, which quickly became one of the biggest MOOC platforms until today (35 million learners by 2020) [136]. Also, certain for-profit platforms were established, like Coursera (76 million learners by 2020), FutureLearn (14 million learners by 2020) or Udacity (11.5 million learners by 2019) [135], whereby the later was co-founded by Sebastian Thrun. Also, the Hasso Plattner Institute (HPI) started its efforts in 2012 as the first European MOOC provider [87], in Potsdam, Germany. In the context of this platform, which is presented in more detail in the following, the various studies of this thesis are implemented and conducted.

1.4.1 The HPI MOOC Platform

Inspired by the success of the first Stanford and MIT MOOCs, a dedicated team at HPI has set out to develop a European MOOC platform in 2012. The initial goal was to offer courses based on the HPI curriculum to the public before the platform was made available to partners as well. Several innovation key factors in the field of online learning were identified: "the synchronization of learners, the possibility of providing the learning materials a little at a time, supplying various feedback tools for self and external evaluations of learning success and linking with a social platform to enable learners the experience of being part of a social (albeit virtual) learning community" [87].

The first version of the platform was based on a highly customized open-source learning management system (LMS). However, since this technical basis was not designed for the MOOC concept—"partly due to the lack of scaling in respect to the number of course participants" [87]—a new platform has been developed completely from scratch according to the findings of the first conducted courses. This platform is still in use today and is being improved constantly. In addition to the new web platform, native client applications for the mobile operating systems Android and iOS have been developed. Many of the platform's innovations are based on research findings that also originate from the team behind the platform itself. Here, research is conducted in particular in the areas of automated and peer assessment, teamwork and collaborative learning, online proctoring, learning analytics, mobile and seamless learning, game-based learning and gamification, and smart assistants.

Course Format

Courses at the HPI MOOC Platform are divided into sections that usually represent course weeks or form thematic blocks. They usually last two, four, or six weeks. The individual sections consist of different learning items. Here, multiple short video lectures are mostly followed by self-tests. Course weeks often close with a graded assignment and at the end of a course a final exam is provided. In between text items are used, e. g., to explain the weekly structure or to point out and link to other resources. There may also be bonus quizzes and surveys. Quizzes support multiple-choice, multiple-answer, and free-text questions. External interactive tasks (such as programming exercises) or (team) peer assessments [147, 148] with more complex submission formats can also be performed. A discussion forum moderated by the teaching team is available to debate course content. Moreover, collab spaces are provided for further social interactions [146]. Here, learners or teachers can open public or private groups to discuss in a separate forum, edit texts collaboratively, start a video chat, share files, and organize calendar events. After the end of a course, it is usually left open for self-paced learning. Graded assignments are then no longer possible and the forum can only be read, as moderation is no longer guaranteed. All of this only describes the usual course scheme from which one may deviate, depending on the instructional design.

Certificates

Different certificates can be achieved by learners for completing courses, and their availability and thresholds can be configured per course:

Confirmation of Participation (CoP): This ungraded certificate is achieved when a learner has viewed at least a certain number of learning items. The default value is 50%.

Record of Achievement (RoA): This graded certificate is gained when a learner has achieved at least a certain number of points. The default value is 50% of all available points of all graded assignments, such as quizzes, external exercises, or peer assessments.

Qualified Certificate (QC): The requirements are identical to those of the RoA. In addition, the learners have to opt-in for a charged online proctoring of the graded assignments.

Key Performance Indicators and Completion Rates

In addition to the number of course enrollments, MOOCs are often judged by completion rates, i. e., by the number of users who achieved a certificate. The particularly high number of 'drop-outs', i. e., learners who at one point simply leave the course, often made these rates look very poor. However, among experts, it was soon clear that the free access to MOOCs and the high number of lifelong learners lead to a late filtering of users who actually want to complete a course. Therefore these rates are not very appropriate to measure the success of a course. Nevertheless, such key performance indicators (KPIs) are convenient for comparing courses. Hence, we have limited the user groups to calculate more reasonable completion rates.

To calculate the completion rate for Confirmations of Participation, we use the number of users who have viewed at least one learning item by the end of the course, i. e., who have at least attended the course once. We call these users 'shows at end'. Users who never show up in the course are called 'no-shows'. To calculate the completion rate for Records of Achievement, we use the number of users who have viewed at least one learning item by the middle of the course, i. e., they have started to attend the course at a date where it is still possible to achieve enough points for an RoA. We call these users 'shows at middle'. However, these assumptions do not include the intention of the learners, whether they are actually interested in achieving a certificate. So far, this cannot be captured with technical means.

1.4.2 Application Domains of the Platform

After the platform was launched, it was gradually transformed into a white-label software so that interested partners from various organizations can also use it under their own brand name. These different instances of the platform—whose development, maintenance, and hosting are provided at HPI—are therefore used in different application domains, which are presented in the following.

openHPI

The initial in-house platform named openHPI started in 2012 with a course on "In-Memory Data Management" held by Hasso Plattner. More than 13,000 participants from more than 100 countries attended the course [87]. Since then, the platform has been offering free courses in German and English based on the lectures at HPI to the public every year. The courses are usually created and held by professors, their staff, or students. The topics range from academic computer science and applied programming to innovation methods and digital transformation. By the end of 2020, more than 870,000 course enrollments were recorded by more than 250,000 registered learners.

openSAP

In 2013, the German-based software company SAP launched its openSAP platform for Enterprise MOOCs. The primary objective is to enlarge the SAP ecosystem by offering free education and trainings for their employees, partners, and customers about their products and business innovations [107]. The openSAP platform is one of the first Enterprise-based MOOC platforms with over 4,900,000 course enrollments by more than 1.1 million registered learners until the end of 2020. A dedicated team at SAP creates and manages the courses together with content experts. Next to MOOCs, the platform also offers complementary podcasts and microlearning formats.

mooc.house

In order to support organizations and companies that do not require a complete stand-alone platform, e. g., because only a few courses are to be offered, the HPI provides the mooc.house platform since 2015. Here, courses or course channels can be acquired, and the HPI can consult and assist during the creation of such offerings on request. So far this offer is used, e. g., by the German Academy of Science and Engineering (acatech), the Charité – Berlin University of Medicine, the Signavio GmbH, the msg systems AG, and the EU-funded projects BizMOOC and CORSHIP.

OpenWHO

Since 2017, the World Health Organization (WHO) offers free courses on OpenWHO to provide frontline responders with knowledge to contain disease outbreaks and manage health emergencies, which especially due to the COVID-19 pandemic has resulted in an enormous growth of users and courses [159]. By the end of 2020, more than 2.1 million learners have enrolled in courses more than 4.7 million times. The courses often differ from the default MOOC format: they are usually shorter and only self-paced. The learning materials are used not only by WHO staff in the field, but also by Member States personnel, e. g., Ministry of Health officials, other UN and partner organizations, international and national non-governmental organizations, as well as the general public, students, travelers, and others [120].

Further Deployments

During the completion of this thesis, other platforms were already launched or planned, but they are not part of studies within this work. Therefore, they are only mentioned here briefly. Since 2019, teachers are trained at Lernen.cloud to use digital technologies in class. The HPI Schul-Cloud offers the necessary technical prerequisites for the practical implementation of the presented concepts. The application and administration of the HPI Schul-Cloud is also the subject of various free courses on the platform, which are offered as official training.

Furthermore, the AI Campus has launched in 2020. This platform, which was funded by the German Federal Ministry of Education and Research (BMBF), provides courses in different areas of artificial intelligence (AI) for the general public. It has been jointly developed by the Stifterverband, the German Research Centre for Artificial Intelligence (DFKI), the HPI, NEOCOSMO, and the mmb Institut.

1.5 Outline of the Thesis

The remainder of the thesis is organized in four chapters, whereby the main part of the work is thematically divided over the following first three chapters. In Chapter 2, we are extending the architecture of the HPI MOOC Platform with a learning analytics infrastructure and additionally incorporate mobile and web analytics. This enables to capture user interactions from the application's different services and thus generate metrics on the learning behavior and success. Based on these new technical capabilities, a learning analytics dashboard for learners is developed in Chapter 3 and is combined with personalized learning objectives to allow self-evaluation of the learning progress and success with data-driven insights. In Chapter 4, a learning analytics dashboard for teachers is implemented to support them in the facilitation and monitoring of their courses with thousands of learners. Lastly, the thesis is concluded in Chapter 5 by a general discussion of the results reported in all studies throughout this work. Apart from a summary of the findings and contributions, general limitations of the thesis are discussed, and practical implications and future research are outlined. Furthermore, several appendices follow in which additional data is reported, e. g., Appendix B with an overview of all courses that are subject to experiments and evaluations throughout this work.

2 Learning Analytics at the Scale of MOOCs

This chapter describes the design and implementation of a learning analytics infrastructure into the service-oriented architecture of a MOOC platform. In addition, the extension by mobile learning analytics and web analytics methods is discussed. The presented approaches are assessed based on the implemented requirements, a case study, and technical limitations.

2.1 Introduction

As MOOCs are used by thousands of learners, a huge amount of learning process data is generated. With methods from the research field of LA, this data can be utilized to understand and optimize learning and the environments in which it occurs [140]. In order to leverage the tremendous research potential, platform providers and vendors have to establish the means and tools for collecting, processing, analyzing, and accessing the produced data. Especially the massiveness of MOOCs leads to a technical challenge. Therefore, this chapter examines the following research question:

Research Question 1: How can learning analytics be enabled at the scale of MOOCs?

However, technical concepts and insights are rarely published especially for modern micro-service-based application architectures like the one of the HPI MOOC Platform. Due to the distribution of data in different services, there is no central place where analytics data can be stored and retrieved yet. Thus, the following sub-question has to be addressed first, the outcome of which serves as a basis for further work:

Research Question 1.1: How can learning analytics be implemented in a service-oriented MOOC platform?

Next to the web-based access to MOOCs with computers and notebooks, the broad availability of mobile devices has enabled mobile learning for online education [157]. Therefore, mobile technologies offer the opportunity to study the area of mobile learning analytics (MLA). It includes the collection, analysis, and reporting of mobile learners' data [2, 17]. If the data of LA and especially MLA is enriched with contextual information, such as time, location, and network connection, as well as device and sensor data, it leads to the term ubiquitous learning analytics (ULA). Context is a crucial factor for workplace learners, self-directed learners, and lifelong learners to integrate learning sessions in their daily life with a wide range of private and professional activities. It can improve the interaction between learners, their devices, and learning environments. Learning materials and tools can be optimized based on the evaluated contextual information [2, 17]. Also, the HPI MOOC Platform faced these new circumstances and brought its platform to mobile devices with native applications for Android and iOS. The following sub-question is investigated to enable a better understanding of the learning behavior and outcome on mobile and stationary devices:

Research Question 1.2: How can mobile and ubiquitous learning analytics be implemented in a multi-client MOOC platform?

Furthermore, originally intended for e-commerce, web analytics (WA) captures users' interactions and reveals valuable insights about the audience, their activity, and behavior for website operators. Therefore, it has especially gained attention from business corporations, which utilize WA for decision-making processes. Consequently, WA has rapidly evolved and is a common technique today that is widely used and no longer restricted to e-commerce websites only [14]. It is clear that WA and LA are subtypes of the general field of analytics and are thus related to each other. Both methods gather and analyze data about users and their interactions on online platforms to understand the audience and their behavior. This data is eventually utilized to derive actions for optimizations. Even though the underlying objective differs, LA may benefit from integrating WA. While LA is a relatively new and active research field, WA is already sophisticated and well-established. Therefore, by using it for analyzing the behavior of learners one could take advantage of its advanced development. Nevertheless, WA tools have not been profoundly used for this purpose so far [19]. Hence, a third sub-question is formulated:

Research Question 1.3: To what extend can web analytics be used in the context of learning analytics in MOOCs?

In order to examine different aspects of research question 1, the three presented sub-questions are addressed consecutively in Section 2.2, Section 2.3, and Section 2.4—before Section 2.5 summarizes the chapter.

2.2 The Learning Analytics Architecture

To provide LA capabilities for different stakeholders, an extension of the platform's architecture is required to collect and process learner data. Therefore, this section presents the design and implementation of a LA architecture in a service-oriented MOOC platform, which is the subject of research question 1.1. Based on the defined requirements, the approach is then evaluated and design recommendations are derived.

2.2.1 Platform and Requirements

In the following, the technical foundation of the HPI MOOC Platform, its architecture, and design decisions are presented first. Afterward, this conceptual understanding is utilized to define the requirements to implement LA in such a context.

From LMS to SOA

The initial version of the HPI MOOC Platform was based on the open-source LMS Canvas to quickly experiment and test the platform with first courses in 2012, which was a pioneering work in Europe [87]. Based on these first insights, a custom-tailored platform has been developed from scratch which fits better to the paradigm of MOOCs, with thousands of learners in a single course and social activity, as well as a better scalability and performance. Therefore, the current platform has been implemented based on the principles of a service-oriented architecture (SOA) with logically separated functionality in individual services [158] as shown in Figure 2.1. For example, the account service is responsible for managing user accounts and the course service manages all information regarding courses and course enrollments. The services can communicate with each other synchronously through RESTful Hypertext Transfer Protocol (HTTP) interfaces, or asynchronously by publishing events on a shared message queue. Currently, there are three clients available for the platform: a web client served by the web service and two native mobile clients for Android and iOS, which use the platform's application programming interface (API).

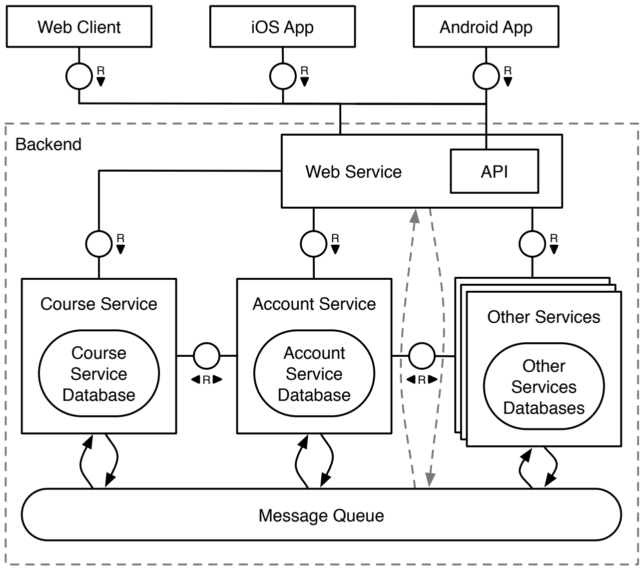


Figure 2.1: The Platform's Architecture

Implementation Requirements

An SOA leads to a distributed data landscape because every service manages its own data persistence layer, and these layers are eventually distributed across different physical machines and rely on different database technologies. This makes it inconvenient when performing analytical tasks. Each service has to offer different analytics endpoints, which can cause heavy load on the overall system and block incoming requests, especially when the data is calculated on-demand. This is due to the fact that microservices are designed to support an operational online transaction processing (OLTP) model. However, the support for online analytical processing (OLAP) is required. In order to overcome this issue, an independent service is mandatory which provides analytics and statistics on separate data stores. Thereby, it has to be extensible to cover different LA use cases of different stakeholders, be flexible to gather data from different system components and clients, avoid high system load and performance impact when gathering and processing data, allow instant data availability, and ensure data privacy.

2.2.2 Related Work

Rabelo et al. [101] showed a big data architecture for LA called SmartLAK that uses an ontology based on the xAPI standard [127]. Data is stored in a Resource Description Framework (RDF) database for high-performance data processing in VLEs to present data insights to teachers in higher education. Since the learning formats of VLEs and MOOCs differ greatly, the findings are only partially applicable in the context of this work. Thus, the use of xAPI is mainly of interest.

Hecking et al. [48] proposed a general LA infrastructure that is extendable and independent of the actual learning environment. Their backend components are designed as an extendable agent system that communicates via a shared workspace using SQLSpaces. They use the ActivityStream format to capture user events and a data warehouse approach for persistent data storage. However, their infrastructure was only evaluated in the context of the Go-Lab environment.

Tabaa et al. [152] implemented a LA system for MOOCs called LASyM based on Apache Hadoop. However, they only used it to detect at-risk learners and it is also not clear why the approach was so strongly coupled to one specific technology.

Ruipérez-Valiente et al. [124] extended the LA capabilities of the Khan Academy MOOC platform with an add-on called ALAS-KA. They use an extract, transform, load (ETL) process to extract the data from the Kahn Academy database, transform it, and load it into their ALAS-KA database. However, their solution is very coupled to the Kahn platform and the Google App Engine environment.

Ruiz et al. [126] also tried to extend the LA support of the Open edX platform, however, this approach seems even more coupled and complex. One of the main issues was the large amount of data, which could no longer be processed in real time. Therefore, they pre-processed the data periodically.

Pérez-Berenguer et al. [100] presented a LA architecture for the INDIeOpen platform, based on two main components: the UPCTforma infrastructure and the INDIeAuthor authoring tool. The infrastructure consists of components for event tracking, event analysis, and learning outcome visualization. Also, they use an interoperability component based on Learning Tools Interoperability (LTI) to connect any learning item from platforms supporting this standard, e. g., Moodle. They use the Caliper standard for event tracking. Events are queued in a message bus and asynchronously processed for cleaning, transforming, and summarizing before they are stored in a MongoDB database. The visualization component is directly embedded to provide predefined dashboards for teachers and students. Based on a custom domain-specific language (DSL), teachers can create learning units directly with the authoring tool. Then, code is generated from the DSL and analytics definitions can be included as well. The approach shown here takes place in a very different context. The actual learning platform, in which the learning units are executed, is seperately from the LA infrastructure, which means that the units have to interact via cross-system interfaces. In addition, the visualizations are not part of the learning platform but the LA components. However, the event-based approach to collect data from external systems can also be applied to our microservice-based architecture. The asynchronous queueing of events and their pre-processing are also promising approaches for high-performance visualizations.

The presented solutions cannot be reused easily and even if a generic LA architecture sounds desirable, conceptual and technical differences of learning platforms make it very difficult to achieve this ideal. Nevertheless, we are able to gain ideas and insights that are incorporated into our approach, e. g., the use of standards such as xAPI, the asynchronous processing of events with a message queue, the ETL process, and especially the independence from a specific database or data processing technology.

2.2.3 Architectural Concept

To implement and fulfill the previously introduced requirements, this subsection explains the concept and architecture of the realized LA service. A complete architecture overview of all system components including the LA service can be seen in Figure 2.2. The service is realized by following the approach of an ETL process, as introduced by Renz et al. [105]. This process is implemented as extensible processing pipelines. Every pipeline consists of an extraction, multiple transforming, and a loading step. The extraction step processes the raw data into a container format. Afterward, the transformation steps process the data and map them to the desired data schema. Lastly, the loading step persists them in different analytics stores. These steps are explained in detail in the following paragraphs.

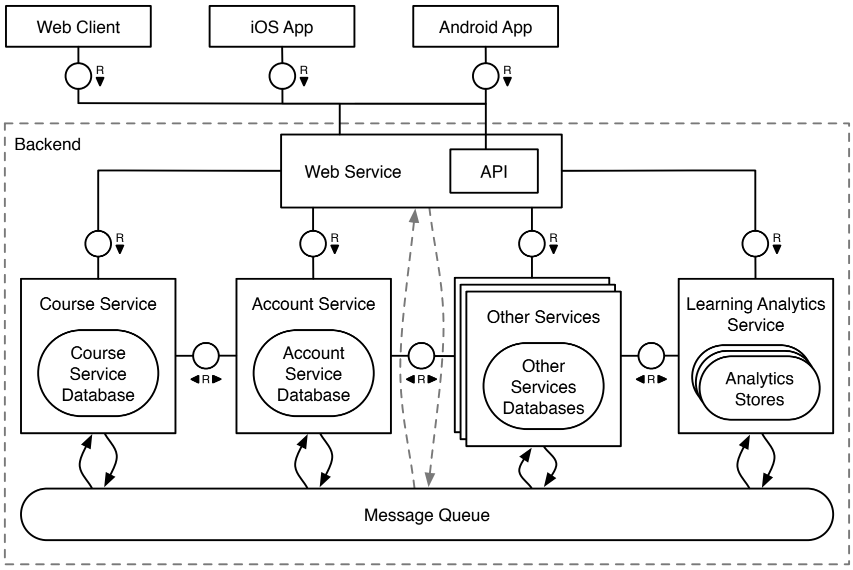


Figure 2.2: The Platform's Architecture with the Learning Analytics Service

Event-Driven Data Collection

The data collection and extraction is implemented by taking advantage of the publish-subscribe message queue. This enables an asynchronous event-driven inter-process communication. Every service can publish events on the message queue. Here, two types of events are used. First, general model changes, e. g., when a model record has been created, updated, or deleted. Second, explicit analytics events, for which a component is provided for the web client to create the events there and transfer them to the backend. The LA service subscribes itself for all analytics events, as well as certain model changes. The queue notifies and passes all corresponding events to the LA service. In this way, the asynchronous non-blocking communication avoids performance impacts on the overall system.

The data structure of the analytics events is inspired by the xAPI [127]: «Actor» does «Verb» on «Object», with «Result» in «Context» at «Timestamp». In the context of the platform, the actor is called user and the object is called resource. The user is the person who triggered the event, the verb is the action performed by the user, the resource is the entity the action was performed and the result is the outcome of the action. The context contains additional information to which the action is related and the timestamp is the moment of the action. Currently tracked events on the platform include—but are not limited to—course enrollments and completions, page visits, quiz and exercise submissions, video player events, download events, forum activities, and helpdesk interactions. An example event is displayed in Listing 2.1 and a complete overview of all events can be seen in Appendix A.

Listing 2.1: Example of an Analytics Event

```json
{
  "user": {
    "resource_uuid": "dfb...578"
  },
  "verb": "VIDEO_PLAY",
  "resource": {
    "resource_uuid": "d84162ce-b711-4d8f-871c-e0660cddd3e5"
  },
  "timestamp": "2020-12-01T15:40:45+00:00",
  "with_result": {},
  "in_context": {
    "current_time": "424.234",
    "current_speed": "1.75"
  }
}
```

Data Transformation with Processing Pipelines

The transformation steps process, enrich, and clean the data. The first step processes the user-agent if the event was sent by the web client in order to identify the user's operating system and browser. The next step determines a coarse location of the request from the IP address to assess the country and city. The third step removes the user-agent and IP address from the event since all crucial information is already extracted from these attributes. They are classified as sensitive personal information, which makes it rather easy to identify a user when anonymized events with hashed user IDs are examined. The last step transforms the data into the appropriate schema of the targeted data storage.

Data Loading into Analytics Stores

The LA service provides the possibility to host different data sources as analytics stores. This provides the advantage to store the same data redundantly—or different data—in various database technologies to optimize query performance. Each data source is configured with its own processing pipeline, whereby the extraction and transformation steps can be reused. The specific loading step stores the data at the end. The general concept of the service and its pipelines is shown in Figure 2.3. Three different pipelines are provided. User interaction events are stored redundantly in an SQL-based data source (PostgreSQL) and in a NoSQL-based data source (Elasticsearch). Additionally, another pipeline is used to enable referrer tracking, which uses the Elasticsearch analytics store as well.

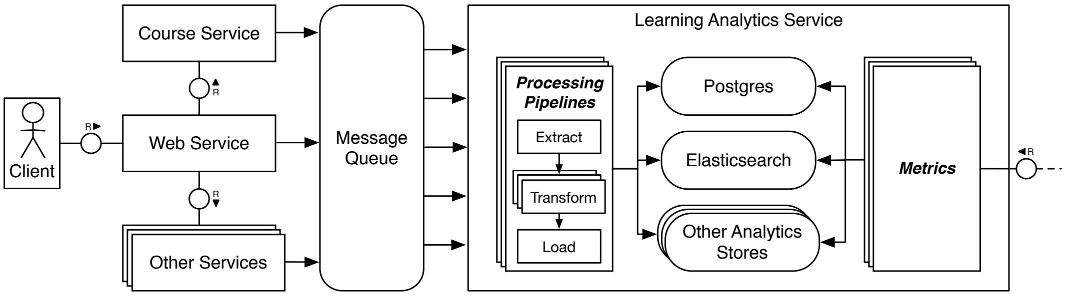


Figure 2.3: Concept of the Learning Analytics Service

Data Analysis with Metrics

Having the data stored in different analytics stores allows to query the data, process them, and expose insights as metrics within the platform. Every metric specifies its data source, optional and required parameters, and a short description with a custom DSL. This enables the provision of a self-documented endpoint for platform developers and researchers to introduce a standardized way to implement new metrics and support the discoverability of available metrics to increase the usage of data-driven insights—either for platform features or research studies. The calculation of a metric provides the possibility of pre- and post-processing of the data, as well as requesting the data source with its native query language. In general, these metrics provide insights about the learning behavior and outcome on a platform, course, user, and user enrollment level.

2.2.4 Evaluation

This subsection evaluates the implemented architecture based on the defined requirements. Therefore, the scalability, extensibility, and versatility are examined. Afterward, the data privacy mechanisms are reviewed. The subsection is closed by a presentation of compiled design recommendations and best practices.

Scalability

Since the implemented approach is used in a real-world MOOC platform with thousands of learners, it must be able to process the incoming data load and provide instant data availability. This means that a user always gets the latest data when requesting a certain metric, which is defined as a processing time for each event of at most one second. To evaluate the data load and availability, we examined a sample period of one year on the largest deployment of the platform at this time. The deployment consisted of four web service nodes and four nodes with all other services, which means that the LA service was also deployed four times redundantly for load balancing. The message queue used to publish events was hosted as a single instance, as well as the PostgreSQL database—which was one of the two analytics data stores. The other data store based on Elasticsearch was operated as a cluster with two nodes.

A total number of 126,180,673 analytics events from 328,507 users was captured in the analyzed period from January 1, 2018, to December 31, 2018, which has resulted in about four events per second on average. Although this number may seem low at first glance, it has to be noted that the general activity on MOOC platforms varies noticeably depending on the time of day, course dates, and deadlines. This results in periods of very high and low activity that have to be considered separately. Therefore, we examined the number of events waiting to be processed in the message queue per hour for the whole year. During the entire period, 67.6% of the time there was not a single event waiting in the queue, i. e., every event was processed instantly. In 31.1% of the captured hour intervals, up to 14,400 events were waiting for a free consumer. This number was chosen since the four LA service consumers were then theoretically stressed with an average of one event per second, which is still considered as instant data availability. Based on this approximation, we achieved a total instant data availability in 98.7% of the time. The higher loads during the rest of the time are probably caused by infrastructure issues and not by activity peaks. To prevent data loss in such outages, all events are stored and kept as unacknowledged in the message queue as long as the analytics stores are unavailable. All in all, we consider our architecture approach proven to be suitable for the scale of a real-world MOOC platform.

Extensibility

An important requirement of the LA service is to provide a flexible architectural design. It has to be avoided to rebuild the whole architecture to include a new schema or data source. Thus, extensibility is ensured with the implemented processing pipeline design. New data, which have to be tracked, can be published by other components through the message queue. Then, the LA service can extract the data within its first pipeline step. A new data source can be added by providing a new load step, which maps the generic event schema to the specific database schema and executes the queries to persist the data. The modularity of the processing pipeline is the most valuable advantage. It can be easily extended or new pipelines can be created by providing additional transform or load steps. Also, every step can be reused by all pipelines.

Versatility

In the following, different use cases and features are explained that are implemented based on the presented LA architecture. This is utilized to assess the versatility of the general approach.

As a typical use case, a teacher dashboard is implemented that visualizes various LA metrics to give an overview of a course. It includes enrollment numbers, active users, and forum activity over time, as well as statistics about learning item visits, quiz performances, geographical learner locations, age distributions, used devices, and learning times. Among other things, it supports teaching teams in identifying anomalies and patterns in their courses such as learning content that is too difficult. Additionally, a learner dashboard is implemented and tested that gives students insights about their own learning behavior. It is based on a concept to better support self-regulated learning. Both dashboards are addressed in this thesis and presented in the next chapters.

Unexperienced teaching teams or limited production times can lead to qualitative weaknesses in MOOCs. Therefore, it is valuable to assist with an automated quality assurance, which LA can enable. Such a concept was implemented by translating best practices into machine-executable rules [106]. These rules are checked periodically and a warning is issued if they are violated, whereby every warning is prioritized and provided with a recommendation for action. Two examples of such rules are quizzes that are too difficult or anomalies in student's video watching behavior such as too many rewinds. Another implemented feature enabled by LA is the cluster viewer, which supports teachers to interactively explore meaningful subgroups of students based on their learning activity to take informed action and measure the effect of executed interventions [154]. Lasty, the platform supports A/B/n testing. With that, researchers can examine new features and compare the learning behavior and outcome of test groups. This is based on LA metrics, which are visualized and compared by their statistical differences and effect sizes [104].

The different presented use cases confirm the versatility of the implemented architecture. It allows realizing a broad range of techniques, ranging from simpler statistics and visualizations to more complex topics like data mining through clustering. Also, various stakeholders take advantage of the LA capabilities, such as teachers, learners, and researchers. This ensures the implementation of further requirements and use cases in the future.

Data Privacy

As the platform is developed and hosted in Germany, the European Union's General Data Protection Regulation (GDPR) is the law in force for governing processing of personal data. Since the LA capabilities are exclusively used to improve the learning experience and optimize the platform and its features, the data processing is considered a legitimate interest. Therefore, no explicit consent is required from the user—as it would be for marketing purposes for example. Additionally, anonymization techniques are applied to further improve the data privacy of the tracked interaction data. Some attributes are omitted which are classified as personally identifiable information, e. g., the user's IP address and the browser's user-agent. No profile data is captured, like the user's name, email, or date of birth. If some data are exported from the platform, the user IDs are additionally obfuscated. To ensure data reduction and data economy, only relevant interaction events are captured, instead of tracking every single click on the platform.

Design Recommendations and Best Practices

Based on the experiences and insights we gathered in more than five years of running the LA service in production on several platform deployments, we have compiled a number of design recommendations for platform vendors and researchers. These best practices aim to support their decision-making when implementing LA capabilities into MOOC platforms.

Concurrent Data Collection and Processing: Analytics, in general, can be seen as an extension to the main application. Thus, the performance impact on the overall application caused by additional analytics tasks has to be kept to a minimum. A common technique is to execute such tasks concurrently. This is realized by utilizing an asynchronous message queue for event collection to avoid blocking the sending components. The data processing is handled by a separate service running independently from other system components.

Schema-Agnostic Pipelining: Different data schemas and query requirements fit more or less well to different storage technologies. Therefore, various analytics data are stored eventually in multiple databases. Hence, we propose a pipeline processing architecture. By utilizing an ETL process for this, all data can be processed based on a generic data schema. Only the last load step converts the data into the database-specific format. This enables a schema-agnostic data processing and minimizes technology and vendor lock-ins.

Reusable Pipeline Components: By utilizing the proposed schema-agnostic pipeline architecture, all transformation processing steps become reusable. For example, this allows applying the same anonymization step to all analytics pipelines. This reduces implementation and maintenance efforts by applying the *don't repeat yourself* principle.

Central Interface for Data-Driven Insights: Instead of having each application component providing its own analytics interface, it is reasonable to have a central interface for data-driven insights. This is realized with an index of all available metrics within the LA service. Also, it abstracts the underlying database technology.

Embrace Open Standards: Through the use of open standards, interoperability with other applications and systems can be achieved best. In the domain of LA, the xAPI format is accepted widely. This standard also defines the learning record store (LRS). Thus, an implementation of such an analytics store can be used right away without further data transformations.

Data Protection by Design: By taking data protection into account in every project stage, privacy risks are reduced and trust increased. Users must stay in control of their data and the benefits of capturing and processing personal data have to be communicated beforehand. It must also be ensured at an early stage that legal requirements like GDPR are complied with.

2.2.5 Conclusion

In this work, an architecture was presented on how LA can be implemented in a service-oriented MOOC platform (research question 1.1). Based on the elaborated requirements, an ETL process was proposed to implement extensible processing pipelines within an independent LA service. This approach utilizes an event-driven asynchronous data collection, a schema-agnostic data processing with reusable steps, and different analytics stores for optimized query performance. It has been implemented for the HPI MOOC Platform and deployed for real-world usage. User interaction events are captured to generate data-driven insights about the learning behavior and create platform features to improve the learning experience and success. Afterward, the architecture was evaluated to study its scalability, extensibility, and versatility by discussing various implemented LA use cases for different stakeholders like teachers and learners. Then, data privacy issues and mechanisms were presented also taking the EU GDPR requirements into account. Lastly, six design recommendations—concerning concurrent data collection and processing, schema-agnostic pipelining, reusable pipeline components, centralized data-driven insights, open standards, and data protection—were introduced. These serve as best practices for platform vendors and researchers to support them implementing LA capabilities in MOOCs.

2.3 Mobile and Ubiquitous Learning Analytics

Users are learning increasingly on the go and with multiple devices instead of being tied to a fixed workstation. To better understand, analyze, and support this learning, this section introduces technical enhancements to the LA architecture to implement mobile and ubiquitous learning analytics in a multi-client MOOC platform and therefore assesses research question 1.2. After that, an evaluation is presented based on a case study to investigate if the usage of mobile devices influences the learning behavior and outcome in MOOCs.

2.3.1 Requirements

From a technical perspective, learning analytics requires a client-side implementation for event tracking and a server-side component for event processing and storage. Since the client-side component for the web frontend and the backend component in form of the LA service have been implemented already, only an extension is necessary here. It has not been captured yet on which type of device an analytics event was created. Therefore, such context information has to be appended. In addition, the mobile applications for Android and iOS also require a client-side component to capture events and send them to the backend. Thereby, the usage of mobile devices causes specific demands on the implementation that differ from the ones of the existing LA implementation for the platform's web client.

Unlike regular computers, mobile devices are frequently exposed to network changes, particularly while still being in use. Hence, the MLA implementation has to be aware of network changes and must avoid data losses. This becomes even more important by the circumstance that the mobile applications can be partially used offline. Because of that, the MLA implementation has to keep the data until a network connection is available again, even if the application has been closed in the mean time. An appropriate data persistence layer is needed to fulfill this requirement.

While using the mobile applications, the user must not be distracted and the device's resources must not be strained excessively by the MLA event tracking. Capturing analytics data is not a user-facing feature and therefore must not change or restrict the user's regular workflow. Due to the fact that most network providers limit the bandwidth after a certain amount of used data, the MLA has to avoid putting excessive strain on the data usage during mobile network connections. Instead, the use of Wi-Fi networks is preferred. Nevertheless, the availability of recent data is important to gain up-to-date data insights. A good balance between urgently needed data and mobile data usage has to be considered.

2.3.2 Related Work

Related research about MOOCs and mobile learning focuses on the conceptual compatibility and synergistic characteristics between both formats [166], as well as how mobile technologies can enrich the MOOC concept [113, 137]. In addition, there are early studies on how MOOC content can be used for microlearning approaches on mobile devices [10] and how mobile-assisted seamless learning can be applied [9] to MOOCs. Tabuenca et al. [153] revealed positive effects of utilizing mobile learning analytics for time management skills. However, learning analytics in mobile learning remains a challenging research issue [2, 5].

Regarding related technical solutions, generic event data tracking and user analytics are widely used in mobile applications. Even if the educational domain of MOOCs lead to a more specific use case, the required technical functionality is similar in all available mobile analytics implementations. It exists a whole industry around this topic, with companies offering mobile analytics as their main product. Also, a lot of larger software companies provide well-known mobile analytics services like Google Analytics for Firebase, Yahoo Flurry, Amazon Pinpoint, or Adobe Analytics. These are commercial products and include multiple components and services: (a) multiple SDKs for event tracking in mobile applications mostly for iOS and Android, (b) a hosted backend system that collects and processes all tracked event data, and (c) a web and occasionally mobile application that presents the processed statistics of the gathered data for reporting.

Since the HPI MOOC Platform already provides a custom backend for the tracked event data, namely the learning analytics service, and LA in general requires more specific reporting as their web or mobile applications offer, only the SDKs of these solutions are relevant for the MLA implementation. However, because of the lack of an open standard for tracking generic analytics data and the strong dependency to their own backends, none of these existing solutions are interoperable with each other. This also applies to open source projects like Matomo or Countly. Therefore, the SDKs cannot be used to communicate with our platform's backend. Nevertheless, these products are widely used and tested, for which reason their documentation and specification can be utilized to gain best practices for the MLA implementation.

2.3.3 Architectural Enhancements

This subsection describes how the event tracking is enriched with contextual data and introduced for mobile, to provide a proof of concept implementation outline.

Learning Analytics with Contextual Data

Based on the following proposed context model (Figure 2.4), it is possible to determine the required data that has to be captured. A wide variety of definitions for the terms context and context-awareness are available [1]. For the platform's domain of e-learning and MOOCs, the entity is specified as the learner and its context is defined as related information about its used device and application, as well as its physical state.

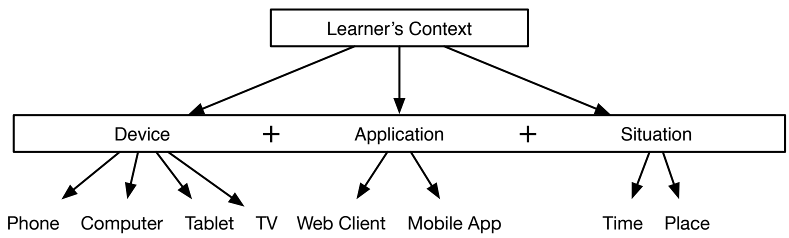


Figure 2.4: Learner's Context Model

To gain these insights, every tracked user interaction event is enriched with additional contextual data as exemplified in Listing 2.2. A basic information about the device is, for example, the installed operating system. This is tracked as the *platform*, on which the application *runs on*. To distinguish if the web client or a native mobile client is used on the platform, the *runtime* is captured, in which the application *runs in*. This can be either a specific browser for the web client or the operating system itself for a native application. To determine the type and size of the device, a device name and data about the screen resolution are captured. Additionally, all events capture the users' time and location. This is a strong indicator about the learners' situation, where and when they use the platform. Additionally, information about the network connection is useful to determine the situation, because a Wi-Fi network is mostly available on stationary and familiar places, like at home or work. A mobile network connection indicates that the learner is on the go or at a foreign place.

Listing 2.2: Example of an Analytics Event with Contextual Data

```
{
  "user": { "resource_uuid": "cc4...6e5" },
  "verb": "VIDEO_PLAY",
  "resource": { "resource_uuid": "78dc954e-c3da-40e2-a678-ea2b93808c6e" },
  "timestamp": "2020-11-30T07:14:02+01:00",
  "with_result": {},
  "in_context": {
      "platform": "Android",
      "platform_version": "11",
      "runtime": "Android",
      "runtime_version": "11",
      "device": "Google Pixel 3",
      "screen_width": "1080",
      "screen_height": "2160",
      "screen_density": "440",
      "network": "wifi",
      "user_location_city": "Nustrow",
      "user_location_country_code": "DE",
      "user_location_time_zone": "Europe/Berlin",
      "user_local_timestamp": "2020-11-30T07:14:03+01:00",
      // [...]
  }
}
```

Mobile Learning Analytics

The core purpose of the mobile learning analytics implementation consists of tracking user interactions in the form of events. These events are sent to the platform's backend through an API endpoint and processed and stored by the LA service similar to the events sent by the web client. Therefore, the same events are implemented for the mobile apps where applicable. Besides, the contextual data are enriched. Since the mobile applications also display content in web views, they have to pass a cookie to the web client with the contextual data. This way, the web client can send analytics events with the contextual data from the actual device when used in the scope of the native apps. Before an event is sent, it is stored in a local database to fulfill the requirement of an appropriate data persistence layer. The SQLite API is used for Android and the Core Data API for iOS. To address the issue of straining mobile data usage by sending events to the backend, every event is tagged with the information whether it will only be transferred in a Wi-Fi network or whether it can also be transferred in a mobile network. This flag is also saved in the event database as well, next to the payload of the event.

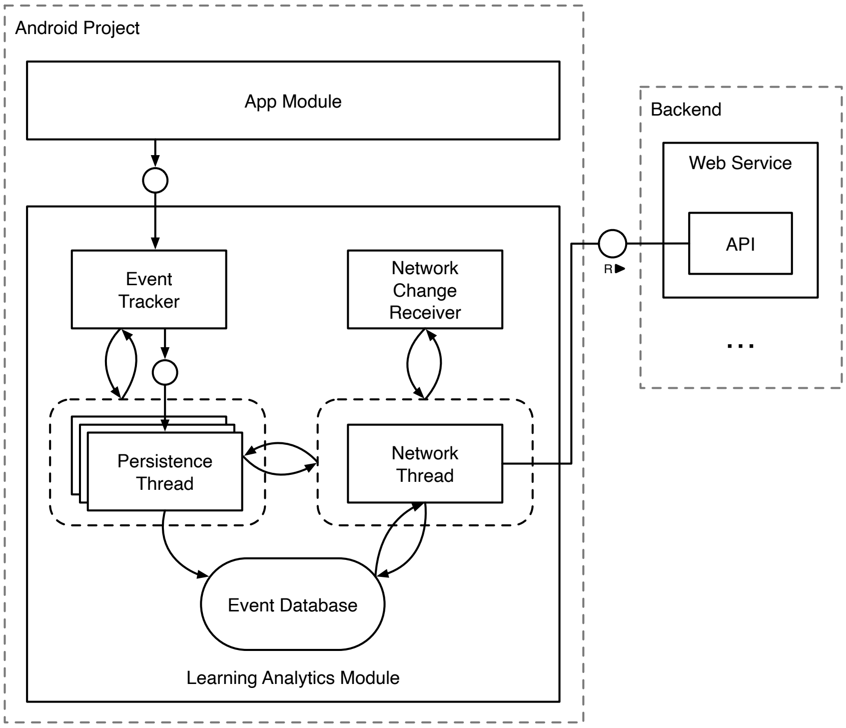


Figure 2.5: Mobile Learning Analytics: Android Architecture

Since the implementations for both mobile applications provide the same MLA capabilities, we only show the Android implementation here as an example. The proposed architecture of the mobile learning analytics module can be seen in Figure 2.5. The main interface of the MLA module is an 'Event Tracker' that receives the user interaction events from the application. Then, the 'Event Tracker' creates a 'Persistence Thread' that updates the 'Event Database' concurrently with the received event, to avoid blocking the main UI thread of the Android application.

Furthermore, a single 'Network Thread' takes care of transferring the persisted events from the 'Event Database' to the backend via an API endpoint. This thread wakes up by a call from a 'Persistence Thread' after the database was updated or by a 'Network Change Receiver' that gets informed when the device changed from an offline state to an active network connection. The 'Network Thread' is also executed concurrently, since Android requires to run network communications alongside the main UI thread, again to keep the application free from long running background work. When the 'Network Thread' successfully transferred an event from the database to the backend, the event gets deleted from the database. This is implemented as a transaction-like communication to ensure that no event is deleted that is not transferred successfully over the network. Additionally, the 'Network Thread' takes care of only transferring events in mobile networks that are not marked as Wi-Fi only to save mobile data usage. Also, events are transferred as batches with up to 50 events at once to save the overhead of creating a HTTP request for every single event.

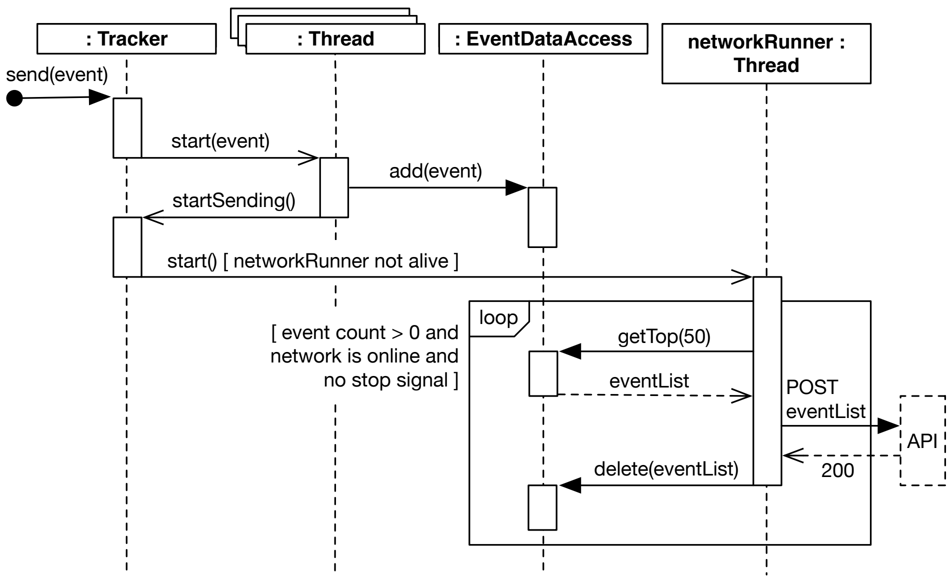


Figure 2.6: Mobile Learning Analytics: Asynchronous Execution on Android

In Figure 2.6 the interaction between the different components is shown. The `Tracker` controls the whole concurrent execution. When a new event is sent to the `Tracker` by the application, it starts a thread asynchronously. The `EventDataAccess` takes care of inserting the event into the database after it was added by the persistence thread. A new persistence thread is started for every event sent to the `Tracker`. The database insert is thread-safe. After inserting the event into the database, the persistence thread calls `startSending()` on the `Tracker`. Also, `startSending()` is called by the `NetworkChangeReceiver` if the application was offline and changes to an online state. This method starts the `networkRunner`, which is also a thread. In contrast to the persistence threads, there is only one network thread at a time. Therefore, `startSending()` only starts a new network thread if it does not already exist and is not running.

The network thread keeps its work in a loop as long as the event count is greater than zero, the network connection is online, and no stop signal has been sent. In every iteration, the network thread fetches up to 50 events from the database. Then, the event list is sent to the backend. If a successful response is received, all events are deleted from the database. The network thread then continues with the next iteration, if the initial conditions are still true, or it stops running. If the network is using a mobile connection, only events without the 'Wi-Fi only' flag are fetched and counted in order to save mobile data usage.

2.3.4 Evaluation

This subsection presents an evaluation of the implemented MLA approach by investigating in a case study whether differences in the learning behavior and outcome can be identified when learning with mobile devices in MOOCs. Therefore, metrics about the user's learning duration, activity, discovery, and performance are examined. Additionally, a new client usage metric is introduced to evaluate the contextual properties of the used devices and applications.

Sample Courses

For the scope of this study four courses are examined, which were running in 2017 and 2018, each two on the openHPI platform and the openSAP platform. All courses had a similar length of six weeks and similar examination modalities: every week was structured into multiple short video lectures followed by ungraded self-tests. At the end of every week, a graded quiz was conducted and at the end of the courses, a final exam was performed. All course contents were completely feasible on mobile to avoid a bias by non-optimized learning items for mobile, such as peer assessments or external exercises. The course topics were "Internet Security for Beginners" (intsec2018[1]), "Big Data Analytics" (bigdata2017[2]), "Enterprise Deep Learning with TensorFlow" (ml2[3]), and "Cloud-Native Development with SAP Cloud Platform" (cp5[4]). Therefore, it can be assumed that the target audience had at least an affinity for IT topics. The courses were certified with a CoP if at least 50% of the course material was completed and with an RoA if more than 50% of the maximum number of points for the sum of all graded assignments was earned.

Methodology

The mobile apps are seen as an additional offering alongside the web platform to enable users to learn anytime, anywhere as a seamless learning approach. Therefore, the learners are divided into two groups: those who used the mobile apps alongside the web platform and those who did not. Both platforms provide an authentic learning environment with real-world users. However, this results in unequal group sizes since it is not a controlled experiment environment and users can decide on their own to use the mobile apps or not.

For each user, different metrics are processed based on the LA events with contextual data. The 'Client Usage' metric is used to devide the learners into the two introduced groups. Additionally, seven metrics about the learning behavior are calculated. The 'Visited Items' metric provides the number of unique visited learning items normalized to the total number of items in a course as percentage, which is the main criterion to gain a CoP. The 'Average Session Duration' is the total duration of all sessions in relation to the total number of sessions. A single session is calculated by all consecutive events with no greater interval than 30 minutes. The 'Quiz Performance' shows the average percentage of correct answers of all quizzes, which is a strong indicator if an RoA was gained. The 'Video Plays' metric shows the percentage of unique watched videos in relation to the total number of videos in a course. Similar, the 'Video Downloads' and 'Slide Downloads' metrics

[1] `https://open.hpi.de/courses/intsec2018/`
[2] `https://open.hpi.de/courses/bigdata2017/`
[3] `https://open.sap.com/courses/ml2/`
[4] `https://open.sap.com/courses/cp5/`

provide the percentage of unique videos and slides downloaded in relation to the total number of videos in a course. Lastly, the 'Forum Activity' shows the sum of all textual forum contributions like questions, comments and answers, as well as forum observations like question subscriptions and question visits, normalized to the number of days between a course start and end date.

For both user groups, all metrics of learning behavior are examined for statistically significant differences using a Mann-Whitney U test for two independent samples. Also, the effect sizes are calculated with Cohen's d for groups with different sample sizes. In Table 2.1 descriptive statistics about both groups' age, gender, and learning outcome are presented.

Table 2.1: Demographics and Learning Outcome of Learners with and without Mobile App Usage

			Demographics		Outcome	
Course	Group	N	Age Mean	Female Quota	CoP Quota	RoA Quota
intsec2018	without app	4162	45.6	0.138	0.281	0.172
	with app	947	39.9	0.154	0.369	0.270
bigdata2017	without app	5805	43.0	0.162	0.279	0.115
	with app	1331	41.2	0.129	0.331	0.184
ml2	without app	8567	39.1	0.090	0.288	0.164
	with app	516	36.5	0.036	0.368	0.225
cp5	without app	4023	39.9	0.114	0.203	0.130
	with app	336	37.3	0.067	0.318	0.250

Results and Discussion

For the learning behavior metrics in Table 2.2, it can be seen that learners who used the mobile apps visited more learning items on average, with a highly significant difference ($p < 0.001$) in three courses and a significant difference in one course ($p = 0.016$). Also, a small effect was measured in three courses ($d > 0.2$). A significant difference was found in two courses for the average session duration, but no practical effect was proven. The quiz performance shows a higher average for learners who used the mobile apps, with a highly significant difference in three courses ($p < 0.001$) with a small effect size ($d > 0.2$) and a significant difference in one course ($p = 0.035$) but without a practical effect size. For all courses, the video plays and video downloads metric show higher averages with highly significant differences ($p < 0.001$). A small practical effect ($d > 0.2$) was identified for the video plays metric of all four courses. For the video downloads metric, two courses had a small effect size ($d > 0.2$) and one course even had an intermediate effect size ($d = 0.708$). The slide downloads metric shows slightly higher averages for users who used the mobile apps with highly significant differences ($p < 0.001$) in three courses, but only one course with a small practical effect ($d = 0.49$). For the forum activity, only one course shows a highly significant difference ($p = 0.001$), but no course had a practical effect.

It can be summarized that users who used the mobile apps visited more items, performed better in quizzes, and watched and downloaded more videos. Highly significant differences and small statistical effect sizes in learning behavior and outcome were identified when learning with mobile devices in MOOCs. However, no significant differences and effects were shown for the average session durations, slide downloads, and forum activities. The demographical means in Table 2.1 regarding age and gender of both groups showed no practical relevance. Nevertheless, the learning outcome based on gained certificates improves on average for learners who also used the mobile apps, which is supported by the findings of the visited items and quiz performance metrics. All in all, the results of this study show that mobile learners tend to be more engaged with the learning material and be more successful in general. However, the causality needs to be examined with further studies.

Table 2.2: Statistics for Learning Behavior Metrics of Learners with and without Mobile App Usage

		Without Mobile App			With Mobile App			Mann–Whitney U		
Metric	Course	N	Mean	Std.Dev.	N	Mean	Std.Dev.	U	p-value	Cohen's d
Visited Items (Percentage)	intsec2018	4162	0.310	0.374	947	0.402	0.407	1675063.0	<0.001	0.243
	bigdata2017	5805	0.326	0.356	1331	0.397	0.380	3323282.5	<0.001	0.195
	ml2	8567	0.343	0.348	516	0.429	0.354	1825712.0	<0.001	0.247
	cp5	4023	0.260	0.326	336	0.343	0.381	622368.0	0.016	0.253
Avg. Session Duration (Seconds)	intsec2018	4162	977.773	1100.513	947	953.391	875.269	1868585.0	0.013	0.023
	bigdata2017	5795	973.816	948.304	1331	909.371	705.853	3797674.0	0.334	0.071
	ml2	8564	817.083	868.271	516	786.128	578.259	2063573.5	0.011	0.036
	cp5	4023	802.863	831.751	336	771.420	716.840	678590.0	0.902	0.038
Quiz Performance (Percentage)	intsec2018	4162	0.341	0.396	947	0.461	0.404	1662879.0	<0.001	0.301
	bigdata2017	5795	0.368	0.369	1331	0.457	0.357	3395540.0	<0.001	0.243
	ml2	8564	0.479	0.412	516	0.598	0.376	1898503.0	<0.001	0.290
	cp5	4023	0.349	0.385	336	0.403	0.373	632467.0	0.035	0.139
Video Plays (Percentage)	intsec2018	4162	0.185	0.338	947	0.315	0.415	1442882.0	<0.001	0.369
	bigdata2017	5795	0.196	0.296	1331	0.281	0.336	3113444.0	<0.001	0.278
	ml2	8564	0.206	0.278	516	0.300	0.295	1629490.0	<0.001	0.335
	cp5	4023	0.182	0.282	336	0.270	0.342	537533.0	<0.001	0.308
Video Downloads (Percentage)	intsec2018	4162	0.090	0.267	947	0.119	0.300	1833878.5	<0.001	0.106
	bigdata2017	5795	0.114	0.264	1331	0.169	0.295	3315855.0	<0.001	0.204
	ml2	8564	0.046	0.163	516	0.167	0.269	1504152.0	<0.001	0.708
	cp5	4023	0.036	0.156	336	0.075	0.214	610122.5	<0.001	0.237
Slide Downloads (Percentage)	intsec2018	4162	0.092	0.254	947	0.123	0.295	1852733.5	<0.001	0.121
	bigdata2017	5795	0.087	0.207	1331	0.118	0.231	3500232.5	<0.001	0.149
	ml2	8564	0.042	0.143	516	0.115	0.233	1762436.5	<0.001	0.490
	cp5	4023	0.044	0.145	336	0.071	0.193	642374.0	0.054	0.180
Forum Activity (per Day)	intsec2018	4162	0.174	1.138	947	0.256	1.856	1980503.0	0.789	0.063
	bigdata2017	5805	0.306	1.540	1331	0.438	2.286	3814122.5	0.441	0.078
	ml2	8567	0.116	0.758	516	0.126	0.409	2052587.0	0.001	0.014
	cp5	4023	0.074	0.493	336	0.093	0.442	650976.0	0.168	0.039

2.3.5 Conclusion

This section introduced an approach how mobile and ubiquitous learning analytics can be implemented in a multi-client MOOC platform (research question 1.2). Therefore, a proof of concept implementation outline was presented that enriched the LA event tracking capabilities of the HPI MOOC Platform with contextual data and introduced an MLA architecture for the mobile applications with a focus on data persistence and network availability. Based on the defined and implemented context model, the tracked interaction events of users' learning activities were processed for different learning behavior metrics, which were examined concerning statistically significant differences and effect sizes between users who only used the web platform and users who also used the mobile applications next to it.

Four courses from two real-world MOOC platforms were studied. It was found that users who additionally learned with the mobile apps visited more items, performed better in quizzes, and watched and downloaded more videos, which resulted in a relevant increase on average course completions. In general, significant differences in learning behavior and outcome were identified when learning with mobile devices in MOOCs. This shows that the MLA implementation extends the existing LA architecture in a meaningful way and new data-driven insights can be gained that take the learners' context into account.

2.4 Web Analytics in the Context of Learning Analytics

Web analytics is commonly used to obtain key information about users and their behavior on websites. Although the foundation of both methods is similar, WA has not been used profoundly for LA purposes. However, large-scale online learning environments in particular can benefit from WA, as it is more sophisticated and established compared to LA. Therefore, this section aims to examine to what extent WA can be utilized in this context, without compromising the learners' data privacy (research question 1.3). For this purpose, Google Analytics is integrated into the HPI MOOC Platform as a proof of concept and limitations are discussed.

2.4.1 Related Work

In general, WA is widely used on the Internet. However, there has not been much research in making use of WA capabilities for analyzing learner's behavior on e-learning platforms so far. Previous work related to this topic is reviewed and presented in this subsection.

Cooper [19] claims that the reasons for the missing utilization of WA tools in the e-learning context are mainly privacy concerns regarding collected activity data. As the majority of WA tools store behavioral data on external servers, control over captured data is lost. He assumes that the majority of e-learning platforms do not utilize services such as Google Analytics, because of their duty of care in handling personal data. In contrast, open-source alternatives such as Matomo, enable operators to control the collected data. However, according to the author, tracking these tools is usually less fine-grained. In general, WA does not meet all needs of LA as it does not cover all information of a learning process that can be useful.

Moissa et al. [91] developed a visualization tool for behavioral data collected in the e-learning environment AdaptWeb that uses Piwik (now Matomo) to capture and store analytics events. Besides the WA tool, the implemented application also retrieves data from the existing database of the platform. The tool provides 20 metrics by combining both data sources, which are mainly based on the number of visits of different types of pages, the frequency of access, used technology, and the utilization of the internal search engine. However, the paper does not reveal, which metrics are based on Piwik and which are computed by querying the local database. In addition, evaluation and limitations of the use of WA in the e-learning context are not discussed as well.

Romanowski et al. [121] integrated Google Analytics into the website of a course of the Penn State University to understand how students interact with it. Page tagging was used for data collection. Results were downloaded from Google Analytics and manually analyzed using Microsoft Excel. Different pages and contents were compared in regard to the number of page views and the average time on a page to discover which features of the website are most effective. The authors concluded that Google Analytics can gather enough data to understand learners' behavior, but has to be combined with further log data of the platform itself to accomplish comprehensive analysis results.

Luo et al. [81] conducted a case study to ascertain potentials and limitations of utilizing Google Analytics for LA purposes in the context of advanced degree online programs. Students' activities of an online course of the Pennsylvania State University were captured using page tagging. For analysis, the researchers considered learner demographics, traffic metrics, efforts of learners, sequence of interactions with contents, and used technology. According to the authors, Google Analytics is well suited for providing an overview of learning processes on e-learning platforms. However, it cannot be used to generate personalized learning reports. Therefore, they inferred that using Google Analytics alone could be too limiting.

In contrast to the assumption of Cooper [19], several big MOOC platforms have integrated Google Analytics in their websites. A manual examination revealed that edX, Coursera, and Udacity have included the Google Analytics page-tagging snippet in their website. EdX specifies in their developer's guide that Google Analytics is used to track all page views and obtain metrics, such as referrers and search terms, used to find the website [33]. Consequently, WA is not used for improving the learning experience of users, but to measure and increase awareness of the platform. However, the other providers do not state their actual intentions and purposes for using WA.

It can be summarized that an integration of WA tools for LA purposes was done only in a basic scope so far. Related work is limited to collecting behavioral data using page tagging and analyzing a fundamental choice of different dimensions and metrics. Although privacy concerns of page tagging are discussed, other data collection methods have not been considered in this context yet. In addition, utilization of more advanced features, such as e-commerce analysis, were not taken into account as well. In contrast, usually the web frontend of the corresponding WA tool is used to manually gather basic analysis results. Hence, a deep integration into any e-learning platform has not been implemented yet. Some limitations of using WA in the context of e-learning were ascertained. Using WA alone could be too limiting to analyze learners' behavior in its entirety. Instead, it can be used in combination with additional LA capabilities to achieve comprehensible results.

In comparison with related research, this work considers the full potential of WA by taking different tools, data collection methods, and analysis capabilities into account. Consequently, results are more meaningful and universal. However, limitations identified by the presented papers are likely valid for this approach as well.

2.4.2 Privacy Concerns

When analyzing user activity, a huge amount of data about users and their behavior is collected and stored. Therefore, privacy laws must be considered when integrating LA or WA into a website [30]. Applicable regulations depend on the type of data that is processed. When collecting only anonymous data, information about individual users cannot be derived and their privacy is not affected. Thus, data privacy laws are only relevant if collected data contains personally identifiable information (PII). As the utilization of WA tools is evaluated using the example of the HPI MOOC Platform, only applicable regulations are examined in the following. Since the service is based in Germany, the European Union's GDPR is the law in force for governing processing of personal data. Art. 4 GDPR defines personal data as "any information relating to an identified or identifiable natural person ('data subject')" and an identifiable natural person as "one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier, or to one or more factors specific to the [...] identity of that natural person". Collecting this kind of data is only allowed if any of the prerequisites listed in Art. 6 GDPR is fulfilled. Among others, this can be the explicit consent of the data subject or the necessity of data processing for purposes based on legitimate interests of the controller. As LA is exclusively used for improving the learning experience of users and optimizing the platform, it is considered as a legitimate interest. Therefore, collecting and processing behavioral data for these purposes is allowed and does not require an explicit consent of the learner. This also applies to the envisaged utilization of WA in context of this work, where additional pseudonymization techniques are applied. According to Art. 5 GDPR, the following principles have to be adhered to when processing personal data [156]:

Lawfulness, Fairness, and Transparency: Data shall be processed lawfully, fairly, and in a transparent manner.

Purpose Limitation: Data shall be collected for specified and explicit purposes only and must not be processed for different purposes than the stated ones.

Data Minimisation: The amount of processed data shall be appropriate to the purpose and limited to its necessary.

Accuracy: Stored data shall be correct and it must be ensured that inaccurate data is rectified or deleted.

Storage Limitation: Data shall be stored only as long as it is necessary for the stated purposes. When used for archiving, research, or statistic purposes, personal data may also be stored for longer periods in accordance with Art. 89.

Integrity and Confidentiality: Appropriate security must be ensured including prevention of unauthorized access, accidental loss, destruction, or damage.

Besides, the GDPR encourages transparency by regulating in Art. 13 that a data privacy statement must be publicly available. This includes information about the person in charge, the purpose for processing of personal data, and, where applicable, the fact that personal data is transferred to a third country or international organization. Concerned persons have the right to obtain information about personal data collected, including a copy of the data itself. Besides, data subjects can claim correction and erasure of any personal data concerning themselves. They also have the right to restrict the processing of their data.

When transferring personal data to third countries or international organizations special regulations need to be considered. Data transmission into European Union (EU) foreign countries is especially allowed if the European Commission has rated the level of data privacy of the corresponding country as appropriate. In this case, an additional approval is not necessary. Otherwise, there are several other conditions authorizing the data transfer, such as concluding a contract between the data exporter (i. e., the MOOC platform) and the data importer (i. e., the WA tool) including standard contractual clauses of the EU according to Art. 46 GDPR. Besides, also the explicit consent of the user legitimates transfer of personal data to third countries.

2.4.3 Architectural Enhancements

To evaluate the applicability of WA in context of LA and thus answer research question 1.3, a WA tool is integrated into the HPI MOOC Platform as a proof of concept. This subsection presents the concept and implementation of this integration.

Choice of Web Analytics Service

There is a great number of different analytics suites available that can be used for the analysis of learners' behavior. Even though this work only aims to evaluate the utilization of WA tools for this purpose, there are still many services to choose from. We decided to integrate only one of these as an example and representative for WA tools in general, as their core features are mainly the same. However, there are differences in regard to more specific and advanced analysis capabilities, processing limitations, and pricing models. We evaluate the proprietary tools Google Analytics and Adobe Analytics, and Matomo as an open source alternative.

Google Analytics is well-established as it is the most popular WA tool. Consequently, it is a paragon in its field and therefore well suited for examining the applicability of WA for LA purposes in general. It comes with a wide range of features, which enable evaluation of different aspects of WA. Even though Adobe Analytics still exceeds these analysis capabilities, the majority of additional features are not applicable in the context of e-learning. Furthermore, Adobe Analytics is highly complex and not as well documented as Google Analytics. Therefore, its integration would be more complicated and costly. In contrast to the traditional, self-hosted setup of Matomo, Google Analytics and Adobe Analytics run on cloud-based servers. As a consequence, it does not have to be taken care of deployment and maintenance of the services. Furthermore, the corresponding services are highly performant, which leads to relatively short response times even for more complex computations. Nevertheless, data privacy could be an issue when storing user activity data on external servers, especially when they are located outside the EU.

Besides, using such an API is required for the envisaged data collection concept. While all examined tools support this type of data collection, there are differences in regard to limitations of the particular implementations. Google Analytics can only process so-called 'hits' that are not older than four hours in this way. The other examined services do not have such a temporal restriction. However, if events are sent to Matomo with a delay, all reports must be reprocessed subsequently to include the new data. In contrast, the Data Insertion API of Adobe Analytics cannot process received data correctly unless events of the same user are transferred in the same order as they were triggered. While one can deal with the time constraint of the Google Analytics Measurement Protocol, limitations of the other tools complicate their integration into the platform's architecture. Due to the asynchronous manner of the existing data collection approach in the platform and the fact that there are multiple instances of each microservice running in production, preservation of the sequence of analytics events during processing is not given inherently. Collected events are concurrently processed by different machines, which may result in mutations in the order of hits. This does not only contradict the restrictions of Adobe Analytics, but would also require regular recalculations of Matomo reports, which is an expensive and time-consuming process. To ensure preservation of the events' sequence, synchronization mechanisms would need to be implemented, which would be complex, cause additional network traffic, and thus have a negative effect on the performance of the service.

All in all, Google Analytics is the best suited WA tool for the purpose of this work if the data privacy concerns are addressed. It supports a broad range of functions, is easy to set up, and satisfies the needs and requirements for integration into the existing infrastructure and architecture of the platform. As the general concepts and main features of WA are the same for all related tools, the findings of this work also largely apply to the utilization of WA in the context of LA in general.

Data Collection

There are different data collection methods available in Google Analytics. The most common and easiest one is page tagging, which requires to insert a small JavaScript snippet provided by Google Analytics into each page. This snippet takes care of gathering needed data and sending it to the WA service. Although integration using this technique is simple and effortless, it comes with some issues. Page tagging slightly increases page loading times as another JavaScript file needs to be loaded and executed. Besides, the existing data collection procedure cannot be used as page tagging would incorporate a separate event tracking. Moreover, page tagging cannot be used properly in the native apps, where mobile software development kits (SDKs) would have to be utilized. This would cause code duplications and is vulnerable for inconsistencies between the different clients. Besides these technical issues, there are also privacy concerns in regard to page tagging as control over data that is sent to the service would be lost.

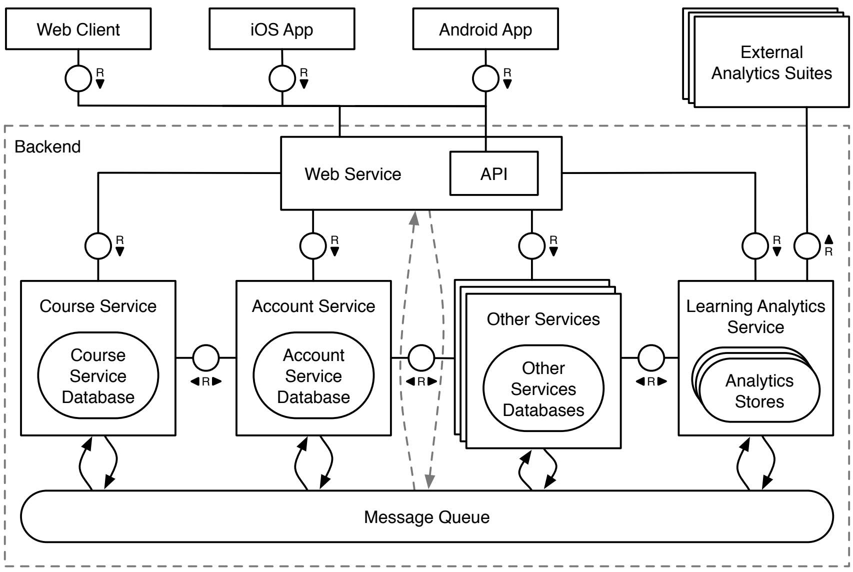


Figure 2.7: The Platform's Architecture with Internal and External Analytics Stores

The HPI MOOC Platform platform already has an analytics infrastructure, which takes care of tracking and persisting certain user activities for LA purposes. In the platform's SOA, the LA service receives interaction events from any client and executes pipelines each representing an independent ETL process. Thanks to the flexible and extensible architecture, we integrate the data collection for Google Analytics into the existing service (Figure 2.7). Thus, a new pipeline is added for transforming interaction events according to the Google Analytics hit schema and emitting them via the Google Analytics Measurement Protocol. The pipeline consists of multiple steps: extraction, enrichment, pseudonymization, schema transformation, and batching before transferring them.

The asynchronous data collection has no impact on the performance of the website. Since tracking is already implemented in the clients, there is no need to adapt either the web client or mobile apps. Instead, all logic related to Google Analytics is encapsulated in the LA service. Moreover, the basis of data stored locally in analytics stores and hits sent to the WA tool are the same, which prevents inconsistencies. As hits are constructed, pseudonymized, and sent manually, we can completely decide, which interaction data is sent to Google Analytics. Thus, control over the data that is sent to third parties is regained.

Mapping Analytics Events to Hits

In order to be processed by Google Analytics, each analytics event must be transformed to a hit, which follows the schema defined by the Measurement Protocol and represents the underlying interaction in the best possible way. Therefore, depending on the event type and available context data appropriate parameters are specified manually based on the different analytics events.

In general, each hit has a type indicating the kind of interaction it describes. Some parameters may be set only for specific types. The types of hits constructed by the mapping are limited to `pageview` and `event`. All events triggered when a user visits a certain page are mapped to `pageview` hits. Otherwise generic `event` hits are created. To ensure data privacy, the SHA-256 hash of the user ID is used, which cannot be used by third parties to identify the user. Also the IP address and User-Agent are omitted, by sending empty payloads. The implementation of this mapping proves that online learning activity can be mapped to WA concepts. However, creating a generic mapping is virtually impossible as each e-learning platform and WA tool has different data schemas and capabilities. This situation can be improved by using a standardized format on the platform side such as xAPI.

Hit Batching and Emitting

The Measurement Protocol supports sending batches of hits inside a single HTTP request. This feature is used in this context to lower the number of requests sent to Google Analytics, and thus increase performance. Batching and emitting hits with an internal message queue prevents data loss and simplifies error handling. As requests are sent to an external service, connection errors are more likely to occur than it is the case when accessing local databases.

One limitation of the Measurement Protocol is that it can only process hits that are not being older than four hours [45]. To cope with this restriction during low activity times, a timeout less than four hours is assigned to each received hit. If this timeout expires before the hit was emitted, all outstanding hits are dispatched even though the maximum batch size has not been reached yet.

For error handling, a message received by the consumer is only acknowledged if it was sent successfully to Google Analytics. Consequently, acknowledgement of messages is outstanding as long as the maximum batch size is not reached. If an error occurs while sending a batch of hits, the corresponding messages are negatively acknowledged. This results in the messages being requeued. Therefore, the consumer receives these messages again, and thus automatically retries sending them to Google Analytics.

Besides, the mentioned data loss issue is prevented. The message queue is configured to be durable, i. e., store unacknowledged messages on disk. This additional persistence layer makes sure, that no hit is lost even if the LA service, the message broker, or the machine running the service is shut down or crashed. The whole process is visualized in Figure 2.8.

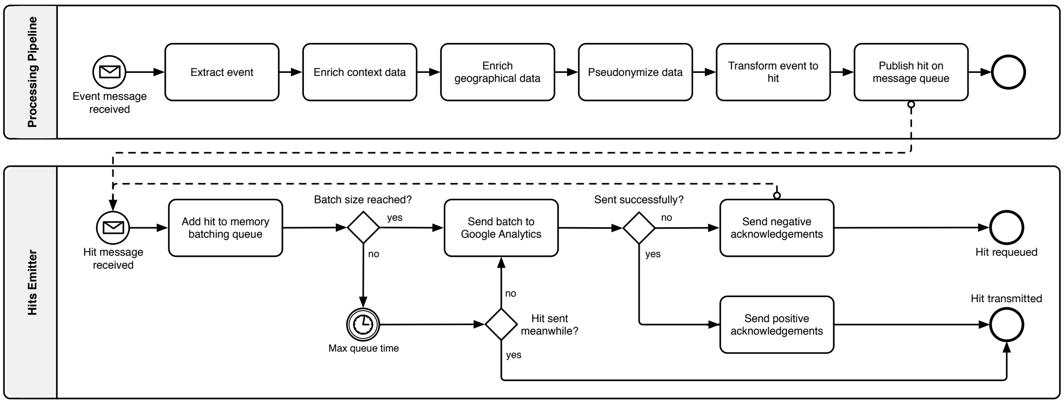


Figure 2.8: Process for Data Collection, Processing, Batching, and Emitting to Google Analytics

Data Privacy

As discussed, data privacy laws must be considered only when processing personally identifiable information. Therefore, it has to be determined first, whether PII is collected. As described before, the hashed user ID is sent to Google Analytics. The ID of a user is considered as a pseudonym of the person and thus as PII according to the GDPR. Even though it is sent as a hash, it is possible for anyone that has access to the platform's databases to identify a certain user given its hashed ID by recomputing all hashes. Nevertheless, Google does not have access to the platform's databases and consequently cannot identify single users. According to the GDPR, location data also belongs to personal data. As the geographical location is retrieved from the user's IP address, it is only a rough estimation of the city or the country of the user's location. Thus, it is not considered as PII.

According to the GDPR, users have the right to receive a copy of collected data about themselves and can claim correction and erasure of these data. Google Analytics provides the possibility to download a file containing all collected data of a certain user. Besides, the entire personal data of a single user can be deleted either in the web frontend or via User Deletion API. On a user's request, the person of authority can take care of providing the copy of data or deleting the data of the submitter. Nevertheless, hits once sent to Google Analytics cannot be modified anymore. Therefore, when a user requests correction of data, they can only be deleted to ensure correctness.

Google is a USA-based company and collected data is stored on servers in the USA as well. Therefore, special regulations could apply, because the USA is a third country from the perspective of the EU, in which the GDPR is not in force. However, Google has a Privacy Shield certificate, which causes the company's level of data privacy to be classified as appropriate to the GDPR. Thus, an explicit approval of the user is not necessary.

Retrieval of Google Analytics Metrics

Two APIs are provided to programmatically retrieve analysis results from Google Analytics. The Reporting API enables retrieval of preprocessed and aggregated reports defined by a certain query. The Realtime API enables retrieval of realtime data. However, this API can only return a small range of basic dimensions and metrics. Google Analytics reports can be used to obtain or derive certain LA metrics, which can subsequently be integrated into the platform. For this purpose, the metrics provided by the LA services are extended by metrics querying Google Analytics.

Instead of replacing existing metrics, this work focuses on integrating new WA metrics into the platform, which are relevant in the e-learning and MOOC context. A typical WA topic which is not well represented in the existing metrics are sessions. The reason for this is that computation of session-related metrics on raw event data is expensive. However, analyzing learners' sessions can help to understand how often and how long users are learning on the platform. Hence, several metrics are implemented querying appropriate Google Analytics dimensions and metrics, such as the average session duration and days since last session.

Another metric that can be easily obtained from Google Analytics is the number of active users at a certain point in time, as it is also common for WA. Therefore, a metric is implemented that returns the number of active learners for each day and hour of a given date range. Besides, another metric is added aggregating this data by calculating the average number of active users for the hours of each day of the week. While the first one can be used to obtain the actual activity of the past, the second metric gives an overview about typical weekdays and daytimes learners are accessing the platform.

Analyzing how users are navigating through a course can help to identify problems of its structure and contents. Unfortunately, it is also an expensive task when working with raw event data only. However, analysis of navigation paths is also a common WA task, which is why two corresponding metrics are integrated. The first one identifies exit items, i. e., items being regularly the last ones within a session and consequently can cause session exits. A high exit rate could indicate that the content is too complex or incomprehensible causing frustration for the learners resulting in session exists. The other metric computes the percentage of back jumps for learning items, i. e., the proportion of page views originating from any succeeding item regarding the structure of the course. If many learners return to a certain previous item again during the progress of a course, it can indicate that prior knowledge being taught in this item was not understood well by a large part of learners.

Limitations

Although Google Analytics can be used to retrieve several metrics that are useful for operators and teaching teams of a MOOC platform, there are some limitations in regard to the kind of data that can be obtained. In general, the web frontend of Google Analytics is used by the majority of customers, which is why Google mainly focuses on implementation of this component. As a result, a few information can only be extracted from the frontend, but not via API. Usually, WA is used to analyze behavior of the entire user base or certain user segments. Therefore, the API does not return any data about individual users. As a result, learner-specific metrics cannot be implemented using Google Analytics. Due to the typical purpose of WA tools, Google Analytics comes with advanced e-commerce analysis capabilities. Among other features, this includes measuring conversions of predefined goals and analyzing the shopping behavior visualized by a funnel that shows at which stages users abandon the buying process. In context of this work, the utilization of these e-commerce features in the MOOC context is examined. For example, the progress of a MOOC can be compared with the purchase of a product. Following this idea, different shopping stages can be mapped to actions concerning a MOOC and vice-versa as shown in Table 2.3.

Table 2.3: Possible Mapping of E-Commerce Steps to MOOC Actions

E-Commerce Step		MOOC Action
Click on a certain product	→	Click on a certain course
View product details	→	View course details
Add product to cart	→	Enroll for course
Remove product from cart	→	Unenroll for course
Several checkout steps	→	Visiting learning items of course
Complete purchase	→	Pass exam and complete course

Based on this mapping, additional hits can be sent to Google Analytics containing the corresponding e-commerce parameters. As a result, respective analysis capabilities can also be utilized. For example, completion rates of courses can be calculated using conversions and corresponding funnels can be analyzed. This makes it possible to identify sections or items of a course that cause learners to abandon the course. The main problem of this idea is the fact that the e-commerce metrics are based on single sessions and cannot be calculated across multiple sessions of the same user. This contradicts the general concept of MOOCs as an entire course can usually not be completed within a single session. Instead, a course typically runs over several weeks and is elaborated by a user in multiple sessions. However, the course progress of a user would not be considered in its entirety, but as several independent attempts to complete the course. Therefore, the e-commerce concept of Google Analytics can only be applied to MOOCs partly.

2.4.4 Discussion

As shown in the previous subsections, WA tools can be successfully integrated into a MOOC platform and thus be utilized for LA purposes. A large part of relevant metrics can be retrieved using WA capabilities. However, there are some limitations, which is why this method cannot be used exclusively in this context. This section examines to what extend WA tools can be utilized to gather insights in learning behavior on e-learning platforms. A great number of LA metrics correspond or can be mapped to WA metrics. For example, the session duration provides information about how long users learn at a stretch and page views indicate how often and in which order learning items are visited. Besides, the number of active users and characteristics of the audience, such as temporal access patterns, used clients and devices, and geographical origins are relevant for both fields. Thanks to generic event tracking, which is supported by the majority of WA tools, any type of interaction can be tracked and therefore analyzed. As a result, especially LA KPIs can be calculated easily using WA as they usually just count the occurrences of a certain event type in a given date range.

As explained before, WA can only cover a part of LA aspects. The main reason for this are conceptual differences between both fields. WA mainly aims to increase traffic and ultimately the revenue of a website. It is all about understanding the audience and generating actions for improving the website itself, but also products and offerings. This includes analyzing users' characteristics and their behavior flow to figure out, which circumstances lead to conversions, such as purchases. Furthermore, evaluating acquisition of new users is a key feature of WA. Besides categorizing users in new and returning, WA tools typically analyze referrer Uniform Resource Locators (URLs) and measure the success of marketing campaigns. In addition to these business aspects, WA tools can also help to identify parts of websites that need improvements regarding content or structure, but also errors and performance issues in general. However, these actions are usually related to the goal of increasing revenue as well, since the user experience and satisfaction have a considerable impact on the buying behavior of customers.

One aspect of LA focuses on optimizing the learning experience on online platforms. This also includes improving the user experience as it is done by WA. However, the intention differs between both fields. While WA aims to help businesses in decision-making processes and is intended to increase revenue, users are in the center of LA as they shall be supported while learning. This also includes encouraging individual learners for example by identifying users at risk that are likely to dropout soon and could therefore need special assistance. This is not a use case of WA and consequently corresponding tools do not typically support analyzing behavior of single users. Instead, only metrics regarding user segments (e. g., mobile users or users of a certain country) or the entire user base can be accessed.

In WA, sessions are the central element for analyses. The amount of metrics that are calculated across subsequent sessions of the same user are very limited in the majority of existing tools. For instance, conversion rates and e-commerce metrics in Google Analytics are restricted to sessions. In contrast, a learning process usually extends over a long period of time. This fact also applies to online learning. For example, the majority of MOOCs has a length of six weeks, in which contents are typically published gradually. Hence, learning takes place in a great number of sessions and perhaps on multiple devices. To analyze the entire learning process of users within a course, activity data has to be considered across sessions. For these purposes, WA cannot be utilized with its current set of features.

To sum up, certain aspects of LA can also be accomplished by utilizing WA. However, the applicability of a metric for being retrieved using WA strongly depends on the type of stakeholder it is intended for. Instances of the HPI MOOC Platform usually have multiple stakeholders: platform owners, teaching teams, learners, and researchers. While researchers are interested in any kind of LA data, the needs of the other three differ. For platform owners, metrics concerning the overall performance of the platform are relevant, which is why they correspond most closely to the typical user of WA. Consequently, the majority of metrics relevant for this role can be queried using WA as well. Especially highly aggregated metrics, such as KPIs, can be easily obtained this way, but are nevertheless essential for platform owners. When it comes to teaching teams, WA can be used only partly. As long as information about the general activity and progress of a course are gathered, it works fine. However, it quickly reaches its limitations when trying to retrieve metrics about smaller groups of users sharing certain characteristics or even single users. As teaching teams are responsible for supervising and supporting learners of a course, WA can help them only with a fraction of their duties. Finally, the method can be used only in a minor extent to provide individual learners with LA data. Metrics about single users, which are most important in this context, cannot usually be obtained. Nevertheless, there are some use cases where WA is helpful.

For example, information about the average performance of learners can help individuals to reflect on their own performance. All in all, WA mainly meets the needs of platform owners, but can assist teaching teams only in some cases. For individual learners there are only rare cases where WA data is relevant.

While this discussion was based on the proof of concept integration of Google Analytics into the HPI MOOC Platform, the key findings also apply to WA and e-learning platforms in general. The revealed potentials and limitations of WA are not specific to the HPI MOOC Platform. This is mainly because characteristics of learning concepts and processes as well as the needs of stakeholders in regard to their use of LA insights are similar among e-learning platforms.

Besides, core concepts and features are the same for the majority of WA tools. Differences exist only in regard to more specific and advanced analysis capabilities, processing limitations, and pricing models. As this discussion considers the general concepts of WA instead of concrete features of Google Analytics, the findings also hold for other tools and are thus valid for the field of WA in general.

2.4.5 Conclusion

So far, WA has not been profoundly used to analyze learners' behavior on e-learning platforms. However, LA may benefit from this sophisticated and well-established method. Therefore, the goal of this work was to examine how WA can be utilized for LA purposes and what limitations it has in this context (research question 1.3).

To answer this question, Google Analytics was integrated into the HPI MOOC Platform as a proof of concept to evaluate the applicability of WA tools in the context of large-scale online learning. For this purpose, the platform's LA service was extended by another processing pipeline that transforms captured interaction data according to the schema defined by Google Analytics and sends it to the WA service. For this transformation, a mapping was developed that models learning activity as WA hits. Instead of using the typical data collection technique of page tagging, the Measurement Protocol is utilized to transfer hits from the platform's backend. Therefore, the solution took advantage of the existing event tracking engine resulting in consistency between the local and external analytics stores. Besides, this approach reinforces data privacy as the amount of data sent to third parties is selected manually.

It was shown that WA can indeed be used to retrieve a large part of metrics relevant in context of LA. However, its applicability highly depends on the type of stakeholder the corresponding metrics are intended for. The needs of platform owners of e-learning platforms and websites in general do not differ much. Hence, the majority of insights relevant for this role can be retrieved using WA. Especially KPIs are essential for this type of stakeholder and can easily be obtained from WA tools. Nevertheless, when it comes to teaching teams, the technique can be utilized only to a limited extent. While WA can provide an overview of the general performance of a course, it reaches its limitations when considering learner-specific metrics since WA is not designed for retrieving user-level information. Consequently, it is also not suitable for providing data insights to individual students. For these purposes, LA-specific methods need to be utilized and WA can be only used to support self-reflection by providing information about the average performance on the platform as a point of reference. Besides, more advanced features of WA tools such as e-commerce analysis cannot be utilized in context of LA due to a mismatch of concepts.

2.5 Summary

In this chapter, the question of how learning analytics can be enabled in MOOCs was explored. Especially the scaling of MOOCs to tens or hundreds of thousands of learners is a technical challenge. To approach this question, three aspects were considered and investigated in the context of the HPI MOOC Platform.

Due to the platform's service-oriented architecture, the data have been separated in service databases so far, which is insufficient for extensive analytics tasks. Therefore, a new LA service was introduced, which had to meet the requirements of extensibility, flexibility, performance, availability, and data privacy. The architecture of the service implements an ETL process based on extensible processing pipelines to handle the analytics data. The data collection of learner interactions is performed asynchronously and event-based. After the transformation through the pipelines, the events are stored in different analytics stores. In this way, various metrics, e. g., about the learning behavior, can be provided and optimized based on the query type and data structure. The scalability, expandability, and versatility of the implemented approach were demonstrated in an evaluation. In addition, anonymization techniques were implemented and the legal context regarding data privacy was discussed. Based on the gained experiences and insights, six design recommendations were derived to support other platform providers and researchers to implement LA in MOOCs. These include: (i) the concurrent data collection and processing, (ii) a schema-agnostic data pipelining, (iii) reusable pipeline components, (iv) a central interface for data-driven insights, (v) the support of open standards, and (vi) data protection by design.

In addition to the web interface, mobile apps for iOS and Android can also be used to learn on the platform. Hence, it is of interest to understand differences in the learning behavior based on the context of the user and to optimize the learning offer based on these insights. Therefore, it was investigated how mobile learning analytics can be supported in a multi-client MOOC platform. For this purpose, two architectural extensions were implemented. First, a context model was developed and technically implemented. The analytics events were enriched with information about the used device, the learning application, and the learning situation, i. e., time and place. Second, the data collection was implemented for the mobile applications as well, which have special requirements. To support situations without network access and offline learning, all analytics data are stored in a local database and only transferred when the network is available. In addition, they are transferred in batches rather than individually, and some event types are transferred only via Wi-Fi to reduce the volume of mobile data traffic. To evaluate the approach, a case study was conducted to investigate whether differences in the learning behavior and outcome can be identified when learning with mobile devices in MOOCs. Seven different metrics about the learning behavior were compared in four courses from two platforms. It was demonstrated that users who additionally learned with mobile applications visited more items, performed better in quizzes, and watched and downloaded more videos. This led to a relevant increase in average course completions. Overall, significant differences in learning behavior and outcome were observed when learning with mobile devices in MOOCs. The implemented MLA approach is therefore a meaningful extension of the previous LA architecture and helps to understand and support mobile learning better.

Since LA is a relatively young field of research, the last aspect was to investigate to what extent methods of the already sophisticated and well-established field of web analytics can be applied in this context, especially concerning the learners' data privacy. After comparing different services, Google Analytics was considered to be the most suitable, due to its large but easily accessible range of functions and integration options. Therefore, selected events were anonymized with a new pipeline, mapped to the corresponding hit schema, and transferred on the server-side. However, some limitations were identified. It is only possible to query aggregated metrics, making it impossible to retrieve learner-specific insights. Therefore, the integration is useful for platform owners, but only partially useful for teachers and researchers, and provides almost no added value for the learners. The mapping of the platform's interaction events to e-commerce processes—the most typical use case of WA—was also proven to be difficult, since conversion rates are usually measured within a session. However, the completion of a course takes place over several sessions. WA can therefore only be utilized to a limited extent in the context of LA, but can be a valuable addition to address specific analytical needs of certain stakeholders.

The elaborated and developed LA infrastructure for a service-oriented MOOC platform serves as a technical foundation for many features and studies. In the following chapters, we discuss how it enables data-driven insights to support learning and teaching in MOOCs. Most of the principles and insights shown can also be transferred to other large-scale learning platforms and thus contribute to the research field of learning analytics architectures.

3 Supporting Learning with Data-Driven Insights

The focus of this chapter is to provide technical support for self-regulated learning (SRL) through data-driven insights. After presenting the pedagogical rationale, a learning analytics dashboard (LAD) for learners is developed and tested in different iterations, and the concept of personalized learning objectives is introduced, integrated into the platform, and coupled with the dashboard. This enables and encourages the diverse and multifaceted learners in MOOCs to set and achieve more individual learning objectives utilizing self-evaluation and strategic planning, which is assessed in several phases with mixed-methods.

3.1 Introduction

Learners in MOOCs have to plan, evaluate, and adapt their learning behavior by themselves continuously to participate in a course successfully and achieve their learning objective since no human teacher can provide individual guidance and support for thousands of learners. This metacognitive skill set is defined as self-regulated learning in the domains of education and psychology and provides several strategies which can be applied by students [99, 129, 174]. However, not all students have these metacognitive skills to reflect on and adjust their own learning behavior.

The research field of learning analytics provides methodologies to understand and improve the learning behavior of students based on their generated learning data [143]. Also, LA offers the possibility for learners to exercise SRL while controlling their learning and decision-making [169]. Anyhow, technical support for SRL in online learning environments is rare and there is an emerging need to further explore LA capabilities to support SRL [162]. We address this gap as we explore the following question in this chapter:

Research Question 2: How can data-driven insights support learning in MOOCs?

There are two main approaches to utilize LA [125]: (1) tools that provide automated recommendations and interventions, and (2) tools which report the data directly to different stakeholders supporting their decision-making with data-driven insights. The latter is often realized with dashboards to visualize data about the learner, their process, and context [134] to help them to reflect and evaluate their learning behavior and outcome [16]. Therefore, learner dashboards are a suitable tool to support and encourage the SRL strategies self-evaluation and strategic planning. To explore this potential benefit for students, we design and evaluate a learner dashboard for the HPI MOOC Platform. Thereby, the following research questions are studied:

Research Question 2.1: Is the learner dashboard accepted and perceived as useful?

Research Question 2.2: Which dashboard visualizations do learners value the most?

Research Question 2.3: Are there differences in the completion rates of learners with regard to the use of the three dashboard variants?

Research Question 2.4: Does the learner dashboard support self-regulated learning?

Since the first evaluations of MOOCs, a main criticism has been the low completion rates ranging from 5 to 13%, which was discussed frequently [25, 61, 62]. This certification-centered focus is reasonable from the perspective of a MOOC platform provider or teaching team since these stakeholders are interested in the success of their courses. Nevertheless, a diverse learning community with different cultural and educational backgrounds comes with many different motivations and intentions [11]. Lifelong learners, especially well-educated professionals, form a large part of the community [34] and a certificate is only one of many different desired outcomes [68, 173]. For example, a dropout can also mean that a learner got all the knowledge they needed at this time [80]. Consequently, the completion-centered perspective of current MOOC platforms excludes a substantial portion of learners. It has to be revised to move beyond the week-based and self-guided one-size-fits-all approach [82] and to better align the students' learning paths and success with their intentions and goals [50, 79].

Unfortunately, courses with self-reported learning goals based on learners' intentions are rarely implemented and conducted. In terms of personalization, the preparation of alternative learning paths, either by varying topics or proficiency levels, requires additional resources. This results mostly in increased production time and cost. Also, modularization can confuse students more than it supports them [63]. Instead, goal-oriented and self-regulated learning have been recognized as a valuable skill set in online learning environments with little support and guidance like MOOCs [69, 173], due to their positive influence on students' achievement [13, 71]. However, the current design of MOOCs neither supports nor motivates learners to complete personalized learning objectives (PLOs) [68] and technical support for SRL in MOOCs is very limited in general [76]. To address this problem domain, this work investigates the following questions:

Research Question 2.5: How successful are learners in achieving their self-reported learning objectives?

Research Question 2.6: How can personalized learning objectives be conceptually supported in MOOCs?

Based on the outlined theoretical concept of personalized learning objectives in MOOCs—to support the SRL strategies goal setting, strategic planning, and self-evaluation—we present a practical study with a focus on the first two strategies by examining the following research questions:

Research Question 2.7: How can personalized learning objectives be integrated into a MOOC platform?

Research Question 2.8: Are personalized learning objectives accepted and perceived as useful by learners?

After that, we integrate the learner dashboard into the concept of personalized learning objectives to enable self-evaluation. Furthermore, we explore the following research questions in order to gain a better understanding which learners select an objective, how satisfied they are with the course, and how successfully they complete their objectives:

Research Question 2.9: How do learners with selected personalized learning objectives differ from the total course population?

Research Question 2.10: Are learners who selected a personalized learning objective more satisfied with the course than those who have not selected an objective?

Research Question 2.11: How successful are learners in achieving their personalized learning objectives?

Before we examine the introduced sub-questions of research question 2, we first explain the pedagogical rationale of self-regulated learning in Section 3.2. Then, the learner dashboard is developed and studied in Section 3.3 (research question 2.1 and 2.2) and Section 3.4 (research question 2.3 and 2.4). Afterward, the concept for personalized learning objectives is elaborated in Section 3.5 (research question 2.5 and 2.6), before it is technically integrated in Section 3.6 (research question 2.7 and 2.8) and finally evaluated in practice when combined with the learner dashboard in Section 3.7 (research question 2.9, 2.10, and 2.11). Lastly, a summary of the chapter is presented in Section 3.8.

3.2 Pedagogical Rationale

Mayes et al. [85] describe learning outcomes of e-learning environments in higher and further education. They extend Goodyear's [44] three kinds of learning in higher education—which are academic, generic competence, and individual reflexivity—by skill-based outcomes to fully encompass further education. They present design principles of learning environments, whereas many researchers recommend to apply constructivism in distance education [53]. They summarize the following principles: (1) the learner actively constructs knowledge, through achieving understanding; (2) learning depends on what we already know, or what we can already do; (3) learning is self-regulated; (4) learning is goal-oriented; and (5) learning is cumulative.

The authors outline two main aspects of activities to construct understanding: interactions with material systems and concepts in the domain, and interactions where learners discuss their developing understanding and competence. In the studied research literature they recognize an increasing focus on the design of learner-centered methods and environments, whereby the ultimate goal of educational technology is the achievement of individualized instruction. Nevertheless, personalization at scale comes with many instructional and technical hurdles. Thereby, goal setting is a first step to understand learners' intention and motivation. Especially in large-scale online learning environments with little support and guidance like MOOCs, self-direction is a critical skill for learners' goal achievement [69, 173], whereas many learners have difficulties in applying self-regulation [75].

3.2.1 Self-Regulated Learning

There is great interest in the factors affecting students' achievement in online learning environments as well as which characteristics distinguish successful from unsuccessful learners. Self-regulated learning has been identified as an important factor positively associated with students' achievement in traditional online learning [13] as well as in MOOCs [36, 71, 76]. It originates from educational and cognitive psychology and refers to the learners' ability to actively and autonomously take control of their learning process [99, 174]. Different definitions of SRL exist while the models of Pintrich [99] and Zimmerman [174] are most prominent. Both describe learning as a proactive and constructive process, wherein learners participate by setting goals, monitoring their progress, and adjusting their learning behavior and actions accordingly, i. e., they show self-corrective behavior. Additionally, they agree that SRL is a skill which can be learned and developed through experience and practice.

Pintrich defines four phases of SRL: (1) forethought, planning, and activation; (2) monitoring; (3) control; and (4) reaction and reflection. For each phase, four different dimensions can be regulated: the cognition (e. g., through activation of prior knowledge and setting goals), motivation and affect (e. g., by building self-efficacy), behavior (e. g., by applying resource management strategies), and the context of learning, i. e., the learning environment. Likewise, Zimmerman describes SRL as a cycle of three phases each encompassing different subprocesses: (1) the forethought phase, including task analysis and self-motivation beliefs; (2) the performance phase, including self-control and self-observation; and (3) the self-reflection phase, including self-judgment and self-reaction.

This shows that SRL is relevant for the preparation, during the actual learning, and in the aftermath of it. Learners should participate in all three phases to be able to successfully regulate their learning.

3.2.2 Metacognitive Learning Strategies for MOOCs

To implement SRL in practice, different strategies are proposed, which can be applied by students. In the context of MOOCs, the metacognitive strategies goal setting, strategic planning, and self-evaluation are of particular relevance [71, 76].

Goal Setting: Setting goals means to agree on a specific goal and the effort that needs to be invested in achieving it. [133, 174]. The goal can then provide guidance for the learning process and serve as a criterion against which the own performance is assessed [99]. Although goal setting is mostly part of the preparation phase, it can be applied in the other phases as well to adjust one's own goals.

Strategic Planning: Strategic planning addresses aspects of selecting proper tasks and how to approach them to eventually achieve a specific goal. For example, learners have to determine the order and timing of activities and select strategies for completing tasks, for instance the procedure and effort invested [99, 174]. Consequently, as part of strategic planning, time and effort management are important strategies.

Self-Evaluation: Self-evaluation requires to set criteria and quality standards against which the learning performance can be assessed, potentially with respect to defined learning objectives [70, 99, 133]. It further implies to monitor the learning progress and outcomes. This enables students to draw conclusions about their learning process and eventually improve their applied learning strategies.

3.3 Towards a Learning Analytics Dashboard for Learners

As introduced, self-regulated learning is a critical metacognitive skill set for students' achievement. However, not every student intuitively self-regulates their learning and therefore technical solutions can help to apply SRL strategies. Dashboards with visualizations about the learning progress and behavior are able to create awareness, encourage self-reflection, and motivate students to plan and adjust their learning behavior. Hence, such LA tools can support the SRL strategies self-evaluation and strategic planning. To explore this potential, a learner dashboard is integrated into the HPI MOOC Platform. This section presents the design process, the concept, and an evaluation of the first learner dashboard iteration. The perceived usefulness and acceptance are investigated (research question 2.1), and in addition, the question regarding which dashboard visualizations are valued most by the learners is examined (research question 2.2).

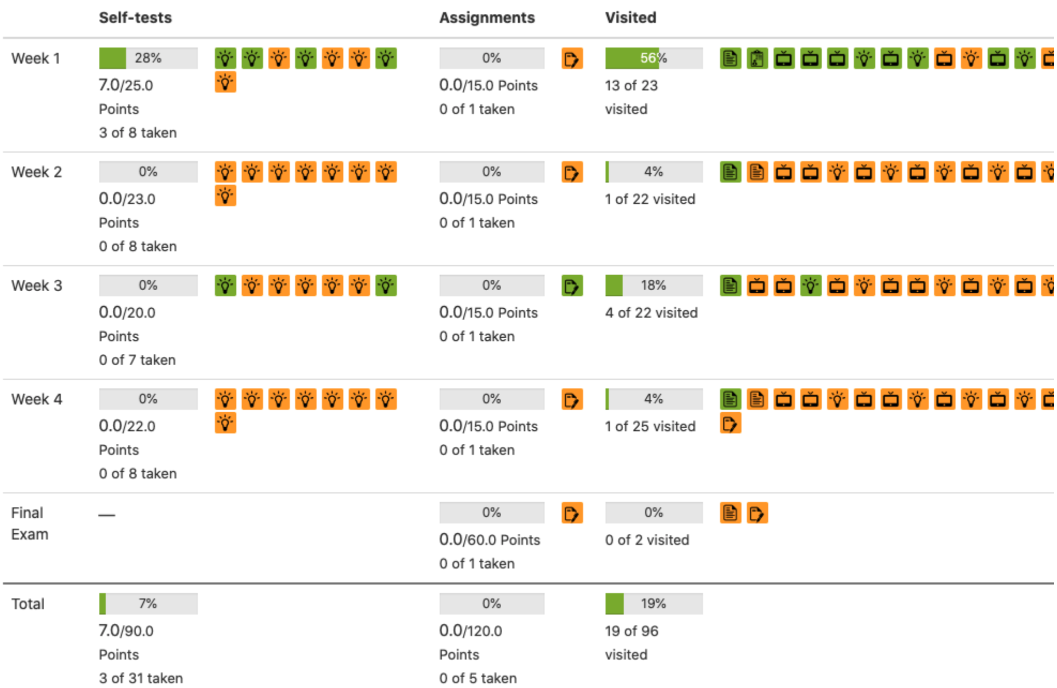


Figure 3.1: The Existing Progress Page of a Sample Course on the HPI MOOC Platform

3.3.1 The Status Quo

To accomplish self-evaluation, the HPI MOOC Platform provides a progress page for each course where students can track their learning progress, which is shown in Figure 3.1. This progress page serves as a starting point to implement the learner dashboard. Currently, it allows the learner to keep track of the overall progression, to get a quick overview of the total achieved points due to the visualization with progress bars, and to identify not visited items. However, the degree to which an individual learning item has been completed is not visible and the items have to be opened consecutively to identify weak points. Furthermore, with a course lasting several weeks and each week containing multiple items, the page gets fraught easily. The most relevant information, which is the percentage of visited items and the achieved points for the entire course, is often not visible at a glance since the page does not fit the screen. Apart from that, there is no further information available that supports students regulating their learning.

3.3.2 Related Work

Since there is a massive number of students, personal support by course instructors is not feasible in large-scale learning environments like MOOCs. Dashboards are a common practice to monitor and evaluate learning progress. They provide a possibility to gain valuable insights into learning behavior and outcomes [161]. However, dashboards in MOOCs so far have been mainly provided for course instructors [77, 126]. Most MOOC platforms therefore still only offer rather general feedback to the students. As an example of the use of feedback systems in MOOCs, Davis et al. [25] created a widget for the edX MOOC platform that allows social comparison with peer learners with the aim of increasing course completion. They found out that the feedback system improves course completion rates, but the benefit of such feedback is limited to highly educated learners. Further, a dashboard for the FutureLearn platform was introduced, which displays demographic information as well as information about a student's learning network, progress, and performance [163]. Beyond full-sized dashboards, the use of smaller widgets throughout the course to provide instant feedback to learners was suggested by several authors [57].

A learning dashboard can be defined as "a single display that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualizations" [134]. According to Verbert et al. [160], effective dashboards are characterized by (1) creating awareness, (2) triggering self-reflection, (3) allowing learners to make sense of the data, and (4) eventually having an impact, i. e., a change in the learning behavior. To move beyond the awareness step, educational concepts have to be the foundation of dashboard design [151]. In order to effectively enable students to make use of dashboards throughout their entire learning process and not just during the actual learning phase, a dashboard has to be adequately integrated into the overall learning design [16]. Jivet et al. [57] further define the following requirements: First, different competences, such as metacognitive, cognitive, emotional, behavioral, or self-regulative competences can be addressed. While most existing dashboards primarily support students on a metacognitive level, only a few consider cognitive and emotional aspects. To adequately support learners, different levels must be approached. Secondly, the information displayed have to be selected deliberately and consider research from the educational sciences. In particular, comparison with peers, which is often intended to motivate students, was found not to be perceived positively by all participants and to even cause contradicting effects [25, 57]. Instead, other reference frames, such as the performance or the attainment of goals, can be preferred to motivate students. Several learning dashboards and visualization approaches exist in general, each focusing on different aspects of the learning process [8, 134]. Based on the experience from designing these dashboards, guidelines and best practices were suggested, such as to outline the learning path to make students aware of how the invested effort translates into outcomes [16]. Last, different evaluation criteria and frameworks for learning dashboards were proposed, which can guide their design [132, 171].

3.3.3 Design Process and Dashboard Concept

To enable self-evaluation, it is crucial to provide students with a visualization of their learning progress, performance, and information about their learning behavior. With this feedback, they can make informed decisions to adapt their personal learning strategies and eventually improve the learning outcome. The currently available progress page on the HPI MOOC Platform offers a good starting point to evaluate one's own current status regarding the course achievement, but does not provide more profound insights into the learning process. For this reason, the platform is extended with a new learner-facing dashboard as the primary feedback tool.

Design Requirements and Challenges

Next to the best practices identified from related work, different requirements respectively challenges have to be considered for the design of the new dashboard: First, the provided information and visualizations must be applicable to all courses and not just to specific courses. Second, due to the weekly release of content on this platform and the often incomplete learning material at the beginning of the course, mainly information about the past can be provided. Nevertheless, the dashboard has to include forward-directed visualizations and recommendations. As the students of the HPI MOOC Platform are used to the current progress overview, the former information must also be available in the new learner dashboard. Third, the dashboard concept has to take an efficient way into account to retrieve the required data since the learners' data are distributed across different services due to the platform's service-oriented architecture.

Design Process

To identify desired features and presumably valuable information, metrics, and visualizations to be incorporated in the new dashboard, an ideation session was conducted. It allowed receiving input primarily targeting the learner's perspective. Seven experts of the HPI MOOC Platform team participated, including teaching team members, researchers, developers, and platform owners. All of them are in close contact with students on a daily basis or learners of the HPI MOOC Platform themselves. After a brainstorming session, the resulting ideas were presented to the group, discussed to clarify questions, and clustered. Subsequently, they were rated by all participants with regard to the perceived usefulness. In the following, the six categories that emerged are summarized and sample ideas are presented.

Progress Overview These ideas focus on providing a better progress overview by collapsing or aggregating less relevant information and information on already completed parts of the course in the dashboard. In contrast to the current basic progress overview, it has to become clearer which learning material has not been finished or fully understood yet. Similarly, the overall progress of the course has to be more prominent, e. g., visible at a glance with a one-color indicator.

Invested Time vs Outcome Making use of a time effort estimation, the dashboard enables comparing the estimated time to the actually invested time, and the aspect of learning efficiency (versus learning performance) can be introduced. For example, the invested effort distributed over the week can be visualized and linked to the learning outcomes, e. g., the achievement of objectives. The given information can also emphasize the most successful learning times.

Time Needed for Attainment While the previous category focuses on the time spent on past learning activities, further ideas target the display of the required time for upcoming tasks, for example completing the remaining course material.

Performance Evaluation As for both the completion of the course and achievement of more fine-granular objectives, the acquired knowledge can be assessed with exercises, for example quizzes. A focus has to be on visualizing aspects of the learners' performance. Not only graded assignments but also self-tests—as a preparation for the exams—can be addressed and compared using various reference frames.

Compare Yourself Additionally, opportunities for social comparison were indicated. This includes the comparison to the course average and students who selected the same objective. It can target the overall course performance, quiz performance, or further aspects, for instance sociodemographic characteristics. Also, objectives or material completed by successful peers can be recommended.

Actions All ideas have in common that the learners are motivated to take action. For example, learners can be encouraged to repeat self-tests, to use the recap mode, or be referred to specific learning content. Besides, discussions in the forum can be fostered by providing students with their forum usage statistics or suggesting forum threads that are relevant. Concerning time management and strategic planning, the next deadlines or steps to complete the course or an objective are valuable information.

Design Concept

Instead of extending the existing progress page with additional metrics, a new learner dashboard is conceptualized and built to address existing weaknesses of the former progress overview. Therefore, findings of the related work, the identified requirements, and the design ideas presented in the previous subsection are considered. However, to entirely integrate feedback for students into their learning process, different widgets can be provided throughout the course to offer instant feedback upon completing activities. In this work, we limit the scope to the new dashboard on a separate page as this is currently the most crucial point for providing feedback to students in the HPI MOOC Platform. The different components of the new dashboard, which address several competences and aspects of learning, are described hereafter. They are grouped by its intended purpose, namely progress monitoring and evaluation, a course cockpit displaying more general course-related information, the analysis of a student's performance, and the provision of learning insights. Also, the concept of utilizing empty states to make recommendations to the learners is briefly described.

Progress Monitoring and Evaluation One of the most important aspects of the learner dashboard is to give an overview of the learning progress. For the general structure of the progress information, the use of a timeline-like visualization is considered first to make it a central part for navigating the course and visually outlining the learners' path through the course. However, due to the issue of limited availability of content and thus often missing material for upcoming weeks, this idea has to be discarded since only the past activities and not the prospective learning path can be depicted. For this reason, the core concept of the former progress overview, mainly visualizing the completed parts of the course, is reused and the key elements known by the learners are adopted. The revised progress overview (Component 1, abbreviated C1) is shown in Figure 3.2.

Since the overall progress summary was often not visible at one sight and the students had to scroll down instead, the most relevant data is moved to the top of the page and visually distinguished from the other presented information using circular progress bars. To reduce the amount of information presented, the most critical performance indicators for each section are aggregated. Progress bars indicate the percentage of completed material and the achieved points for graded respectively ungraded exercises. This is complemented with time effort information in terms of the time remaining to complete the material of a section to enable planning activities.

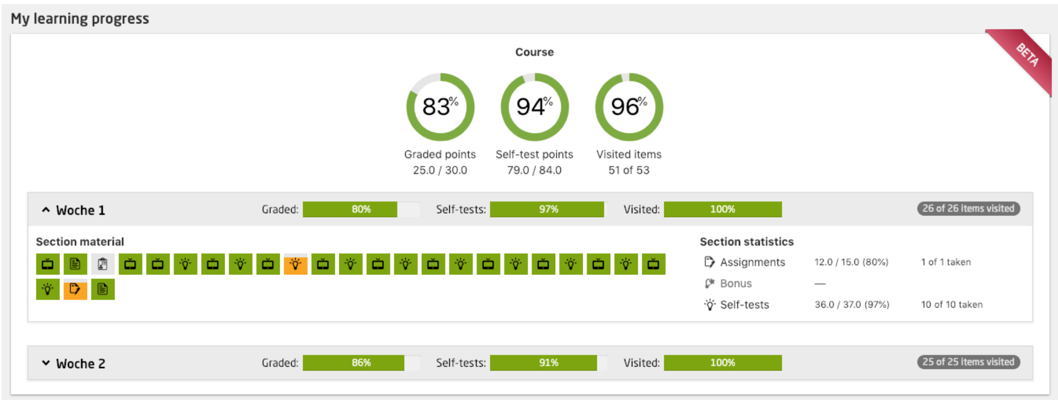


Figure 3.2: The New Progress Overview (C1) on the Learner Dashboard

In contrast to the previous progress page, the details for a section, for example the visited learning items, are not visible by default reducing the space for each section. With that, the progress does not typically exceed the screen size, and it is possible to skim the progress for each section. If required, students can explore more detailed information by expanding the section details. The contained learning material (items visualized as rectangles) for each section is presented similar to the basic progress overview. It is complemented with a time effort estimation and the visualisation of the completion percentage for each element. This introduces different states distinguishing visited and finished content. The degree, to which an item was accomplished, is indicated graphically facilitating students to identify items they may want to review. To offer navigation support, links are provided to directly guide students to the respective content and allowing them to take action immediately. This part primarily targets the metacognitive level by raising awareness and allowing to monitor the progress. Further, it addresses the cognitive level as it aims to support goal achievement and performance improvement [57].

Course Cockpit The course cockpit, shown in Figure 3.3, provides an overview of aspects related to time management and the learners' participation in forum activities. First, the time needed to complete the course (C2) is visualized, which aims to facilitate time management and strategic planning. Additionally, relevant course dates (C3) are shown to raise the learners' awareness of upcoming course events. These primarily include submission deadlines and the release of new learning content. When contributing to the forum by asking or answering questions, a student engages with the content more deeply and reviews different aspects of a topic. To motivate learners to participate actively in discussions, their forum activity for the course (C4) is stated, and it is recommended to use the forum. Consequently, the focus is on metacognitive and on behavioral skills, for example, to motivate learners to explore the forum.

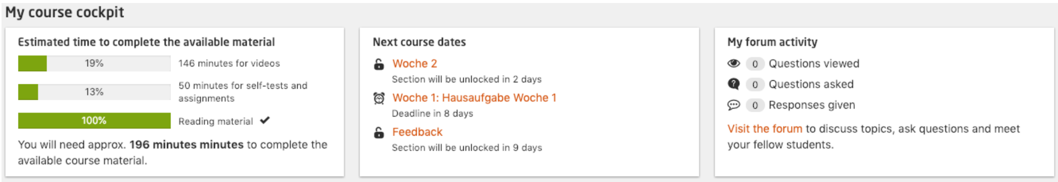


Figure 3.3: The Course Cockpit on the Learner Dashboard. It includes the estimated time for course material (C2; left), next course dates (C3; middle), and the students' forum activity (C4; right).

Performance Evaluation The third part of the dashboard addresses the learners' performance in quizzes and other exercises as shown in Figure 3.4. To support reflection on their learning habits in terms of the exercises approached and to motivate them to complete both the self-tests and the assignments, the performance regarding these types of exercises is contrasted (C5).

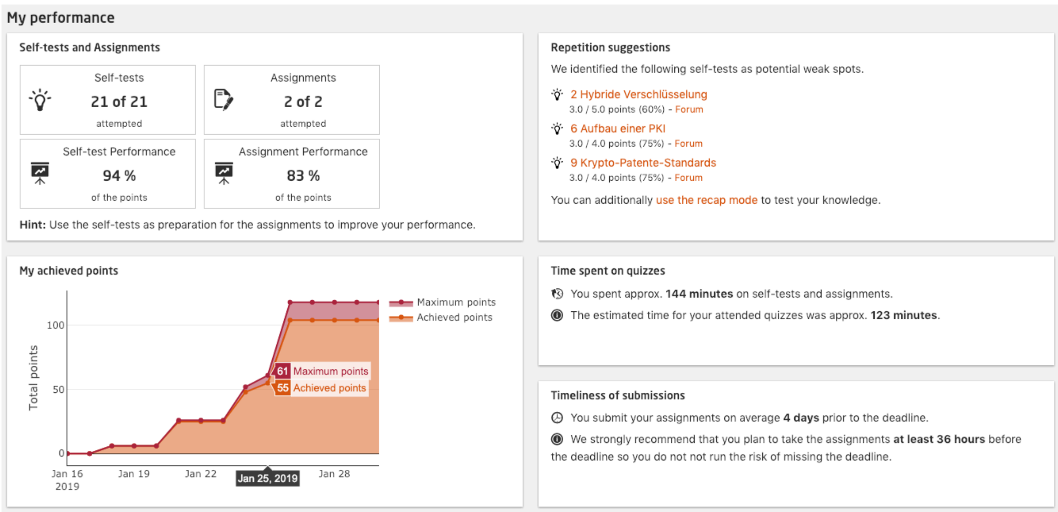


Figure 3.4: The Performance Evaluation on the Learner Dashboard. It includes information on the self-test and assignment performance (C5; top left), repetition suggestions (C6; top right), the achieved points over time (C7; bottom left), the time spent on quizzes (C8; center right), and the timeliness of submissions (C9; bottom right).

In addition to the overall statistics, the self-test exercises with the lowest achieved score are suggested to the learner for repetition (C6), as these may be possible weak spots. Links are provided so that the student can directly start working on the self-tests or ask questions in the forum. Also, the recap mode allowing learners to repeat the questions in an index card manner is suggested since this feature is not often discovered. Besides, a diagram visualizes the accumulated achieved points compared to the maximum possible points of attempted quizzes (C7) to provide feedback on the historical development of the student's performance. Two more metrics target the learning strategy in terms of time management and strategic planning skills. The time spent on quizzes (C8) is compared to the estimated time for the attempted quizzes (self-tests and assignments), which is calculated based on a time effort estimation. With that, students can determine if the time spent on these exercises relates to the learning outcome. If there is a discrepancy, a student may need to better prepare for the quizzes or adapt the applied strategies for learning. Last, the aspect of planning the learning sessions is stressed by showing the timeliness of the submissions (C9) with respect to the submission deadline. When a learner actively plans to work on assignments a certain period before the submission deadline, the risk of missing it due to unexpected personal schedule changes is reduced.

Learning Insights The final section gives students further insights into their learning (Figure 3.5). It is meant to help them analyze and improve their learning process. First, a diagram visualizes a student's activity in terms of items visited over time (C10) to stress the invested efforts towards the attainment of personal goals. While one curve indicates the unique items visited, another curve displays the total number of item visits.

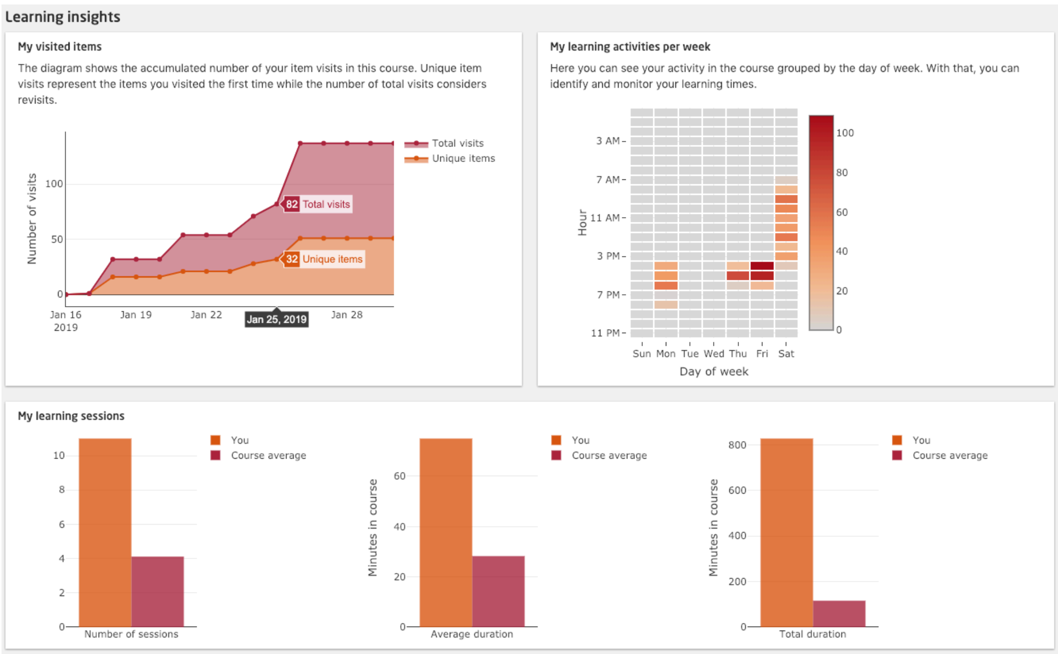


Figure 3.5: The Learning Insights on the Learner Dashboard. These include a visualization of the visited items over time (C10; top left), the learning activity (C11; top right), and session statistics (C12; bottom).

Such visualizations of progress over time are still rare in MOOCs [8]. However, they can help students to identify peaks or patterns of learning time and productivity. A similar approach is offered by the heat map giving an overview of a student's overall course activity distributed over the days of a week (C11). Potential effective learning times can be identified, and the information can be used to allocate times for learning. Last, the students' number of learning sessions respectively average and total session lengths are depicted and compared to the course average (C12). For the other parts of the dashboard, we deliberately avoid comparison with peers as it can have a demotivating influence on learners [57]. The comparison of session time, however, is not critical since students who invest less time can successfully achieve their own goals, too. In addition to the described components, several other metrics can be valuable as well. However, since the dashboard already contains a notable amount of information in this first version, the number of metrics was limited to the presented ones. Too much information at once can overwhelm students and thus rather discourage than support them. After evaluating the effect of the presented information, the dashboard can be extended in subsequent iterations.

Empty States To receive feedback on the submission timeliness or the time spent on quizzes, a learner must have submitted an assignment or self-test. Moreover, self-tests can be suggested for repetition only if the learner has not performed well in at least one quiz. Until then, no data is available, and the display of a dashboard component is empty. A concept applied for the new learner dashboard is the active use of these empty states to encourage learners to reflect and, if necessary, improve their learning strategies. Therefore, hints for learning strategies, motivational statements, or links for tools that can be useful for the learner are provided until data is available for a component. This approach can motivate learners to persist and take action to improve their learning success. The empty state of the repetition suggestions component (C8), i. e., when no suggestions can be made for a learner, is shown in Figure 3.6.

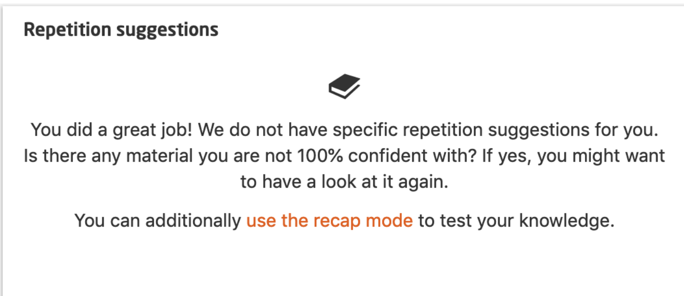


Figure 3.6: An Empty State on the Learner Dashboard. Displayed for the repetition suggestions (C8).

Technical Concept

In contrast to the old progress page, the dashboard is implemented using the single-page application framework Ember[1]. At the time, this was the desired technology for the entire course area for learners. To be able to load all data asynchronously, a dedicated API is provided. It communicates

[1] https://emberjs.com/

with the LA service, among others, where metrics for the new visualizations are implemented. An Ember-Plotly[2] plugin is used to display the data. As the name already indicates, this plugin enables the usage of the well-known Plotly[3] visualization library within Ember. With this, three visualization components are developed in particular: (1) a heat map, (2) a scatter plot, and (3) a bar chart component. The various dashboard visualizations are then implemented as independent data-loading components. Thus, they fetch their data asynchronously to render the visualizations as soon as their data is available. This is especially relevant if some complex metrics have a larger load time and must not block the display of the entire dashboard. For this, a promise object is used that also behaves like an Ember object which updates the template as soon as the data is received. An example for this is displayed in Listing 3.1.

Listing 3.1: Example of an Ember Component to Asynchronously Load Data

```
export default Component.extend({
  pinboardStatsData: computed('course', function() {
    const promise = $.getJSON(`/api/v2/learning_insights/pinboard.json?course_id=${
        this.get('course.id')}`);
    return DS.PromiseObject.create({ promise: promise });
  }),
  pinboardStatsAvailable: bool('pinboardStatsData.isFulfilled'),
});
```

3.3.4 Evaluation

To assess the perceived usefulness of the new dashboard, a survey was carried out in a sample course. It also aimed to better understand which aspects of the dashboard are particularly valuable for the learners. The evaluation of the survey elaborates and answers both research questions 2.1 and 2.2.

Methodology

The survey was sent to students who took part in a course on "Data Security on the Internet" (informationssicherheit2019[4]) on openHPI. The course was part of a series of three consecutive courses on different cyber security topics. This first workshop, for which 4,354 learners were enrolled at course start, was held from January 16, 2019, until January 30, 2019. Since the survey was sent after the end of the course, the results are likely influenced by a certain survivorship bias of the students. In total, 217 learners completed the voluntary survey. The complete response data can be seen in Appendix C.1. The new learner dashboard was enabled for this course at start.

When conducting the survey, the agreement with given statements had to be rated on the basis of a five-point Likert scale. First, six statements targeted the perceived usability and usefulness of the new learner dashboard in general. Subsequently, additional questions examined the value of

[2] https://github.com/EmberMN/ember-cli-plotly/

[3] https://plotly.com/javascript/

[4] https://open.hpi.de/courses/informationssicherheit2019/

the dashboard's individual components. The students' perception of the components is evaluated as follows: the answer possibilities of the Likert scale are associated with a weight ranging from -2 for 'strongly disagree' to $+2$ for 'strongly agree'. Based on the participants' answers, the mean is calculated, resulting in a final rating where $+2$ is the maximum possible score. From this, a ranking of the components is created. Beyond these two parts of the survey, two questions following the "I like, I wish" method aimed to encourage open feedback.

Analysis and Discussion

Two aspects are evaluated with the survey. First, the perceived usefulness and usability is elaborated. Second, the different components of the dashboard are ranked according to their value to learners. To start with, both the usability and usefulness of the dashboard are perceived remarkably well by the participants. Notably, the vast majority agrees or strongly agrees that it is easy to use (94.92%, Q1) and it is quickly apparent how it is operated (88.48%, Q3). It is considered as extremely useful by 84.79% of the learners (Q4), which clearly shows the acceptance of the tool.

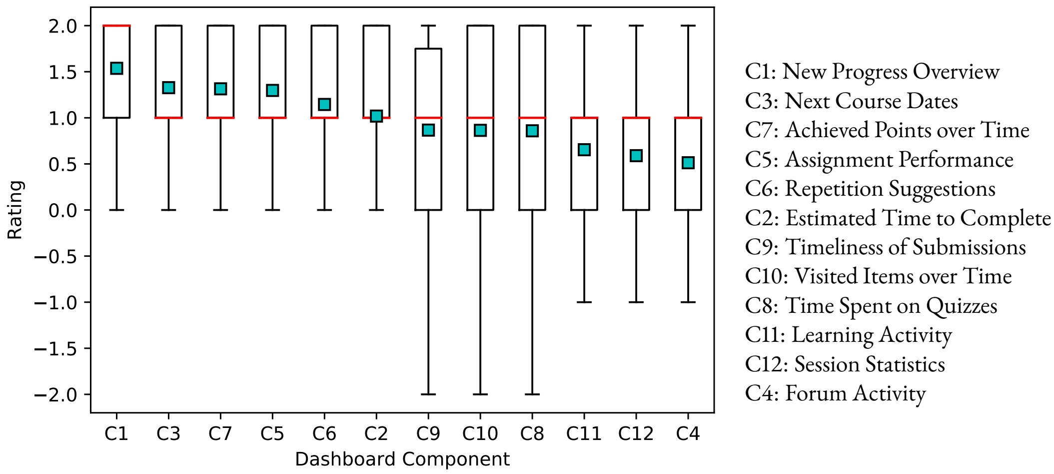


Figure 3.7: The Rating of the Learner Dashboard Components. The red lines are the median values, the blue squares show the mean values.

The learners' rating of the dashboard's components results in the ranking shown in Figure 3.7. Of all components, the students particularly appreciate the progress overview (with a rating of 1.54; C1). This has been expected since providing an overview of the learning progress is the most important aspect of the learner dashboard. Interestingly, it seems that the new dashboard can provide a better and clearer overview of the learning progress for many learners compared to the basic progress overview as this was repeatedly stated. For example, one learner especially valued the circular progress bars as they provide a "quick overview of the percentage of points achieved". Second, the course dates (1.33; C3), the visualization of achieved points over time (1.31; C7), and the separately presented information on the performance in self-tests and graded exercises (1.30; C5) have almost the same rating. The repetition suggestions (1.14; C6) and the time estimation for the course material (1.02; C2) follow next. Fourth, the submission timeliness (0.87; C9), the visualization of the visited items over time (0.86; C10), and the time spent on quizzes (0.86; C8) show a decent rating as well. The items rated lowest by the participants are the heat map for the learning activities (0.65; C11), the session statistics (0.59; C12), and the forum activity overview (0.51; C4). From this ranking, future development directions and improvement possibilities for the dashboard can be derived. To summarize their overall satisfaction with the new dashboard, the students had to rate it with one up to five stars. 87.10% of the participants awarded four or five stars, while the average rating is 4.28 with a standard deviation of 0.89. This shows a high overall appreciation of the new dashboard by the learners.

Besides, qualitative feedback was received. In general, the learners value the new visualizations of the dashboard as they provide valuable insights. Further, the participants like its overall (visual) design. Regarding the components, specifically the progress overview and the course cockpit are mentioned as being useful. These components support planning activities and can be used to navigate the course. Last, participants stress their satisfaction with the choice of information provided and like the clear presentation of the data resulting in a good overview. The latter aspect, however, is perceived ambivalently since several learners also mentioned that the dashboard is overloaded and difficult to comprehend. Mostly, they ask for better explanations for the visualizations presented. Learners further suggest introducing configuration options. This can benefit students who are overwhelmed by the amount of information.

3.3.5 Conclusion

This study introduced the design process and concept of a learner dashboard for the HPI MOOC Platform to encourage self-regulated learning and in particular the strategies self-evaluation and strategic planning. The dashboard provides a general progress overview and insights about the estimated time for course material, the next course dates, the students' forum activity, the quiz performance, repetition suggestions, the achieved points over time, the time spent on quizzes, the timeliness of submissions, the visited items over time, the learning activity, and session statistics. This first iteration of the learner dashboard was evaluated with a survey. In general, the usability and acceptance are perceived very well by the majority of students (research question 2.1). The most valued components are the progress overview, the course dates, the achieved points over time, the quiz performance, and repetition suggestions (research question 2.2). However, improvements were suggested, e. g., more explanations of the displayed visualizations and configuration options to reduce the amount of visible information. These suggestions are considered in the next iteration of the dashboard. Also, we examine whether the dashboard supports in the application of the presented SRL strategies.

3.4 The Learner Dashboard for Self-Regulated Learning

Based on the findings of the last section, in this study we revise the learner dashboard. Afterward, we investigate by means of an A/B/n test if there are differences in the completion rates of learners regarding the use of the three dashboard variants to address research question 2.3. Lastly, we evaluate with a survey whether the dashboard supports learners in applying self-regulated learning (research question 2.4), particularly concerning the strategies of self-evaluation and strategic planning.

3.4.1 The Revised Learner Dashboard

This subsection briefly presents the implemented changes of the LA widgets on the learner dashboard. These are largely based on the concept of the last study (Subsection 3.3.3) and additionally incorporate most of the user feedback. Specifically, the widgets are rearranged within their functional domain based on their perceived value. Besides, further explanations are added to the displayed visualizations and the session statistics are removed entirely. The complete revised widgets on the learner dashboard are shown in Figure 3.8. Since the progress overview is not modified (Figure 3.2), it is not discussed here again. The progress overview is still located on the same page above the widgets.

Course Cockpit

(a) In the course cockpit area, mainly the next course dates are moved to the front, as these were rated most valuable by users after the progress overview.
(b) For the estimated time to complete the available material, a note to include them in the planning of learning times is added.

Course Performance

(c) In the course performance area, the time spend on quizzes and timeliness of submissions are interchanged.
(d) The repetition suggestions are supplemented by the remark to ask for advice in case of uncertainty which is linked to the course forum as a call-to-action.
(e) The achieved points visualization is provided with a detailed explanation: "The diagram visualizes your progress in terms of achieved points of attempted quizzes over the course run time. The achieved points include results from self-tests as well as assignments. Your achieved points are compared to the respective maximum possible points. Use this information to discover your full learning potential!"
(f) Also, the time spend on quizzes is extended with the hint to use the comparison to find out whether the learning content was already internalized and can be recalled quickly, or whether working through the quizzes takes more time.

Learning Activities

(g) In the learning activities area, which was previously called learning insights, the visited items visualization is provided with a detailed explanation: "The diagram shows the accumulated number of your item visits in this course. Unique item visits represent the items you visited the first time while the number of total visits considers revisits. Use this information to include your usual repetitions when planning future learning times!"
(h) Also, the heat map of the learning activities per weekday is improved and an explanation is added as well: "Here you can see your activities in this course based on the day of the week. Identify your usual learning times and include them in your future weekly planning!"
(i) As already mentioned, the session statistics are completely removed. Their perceived value was very low and their computation was imprecise and expensive.

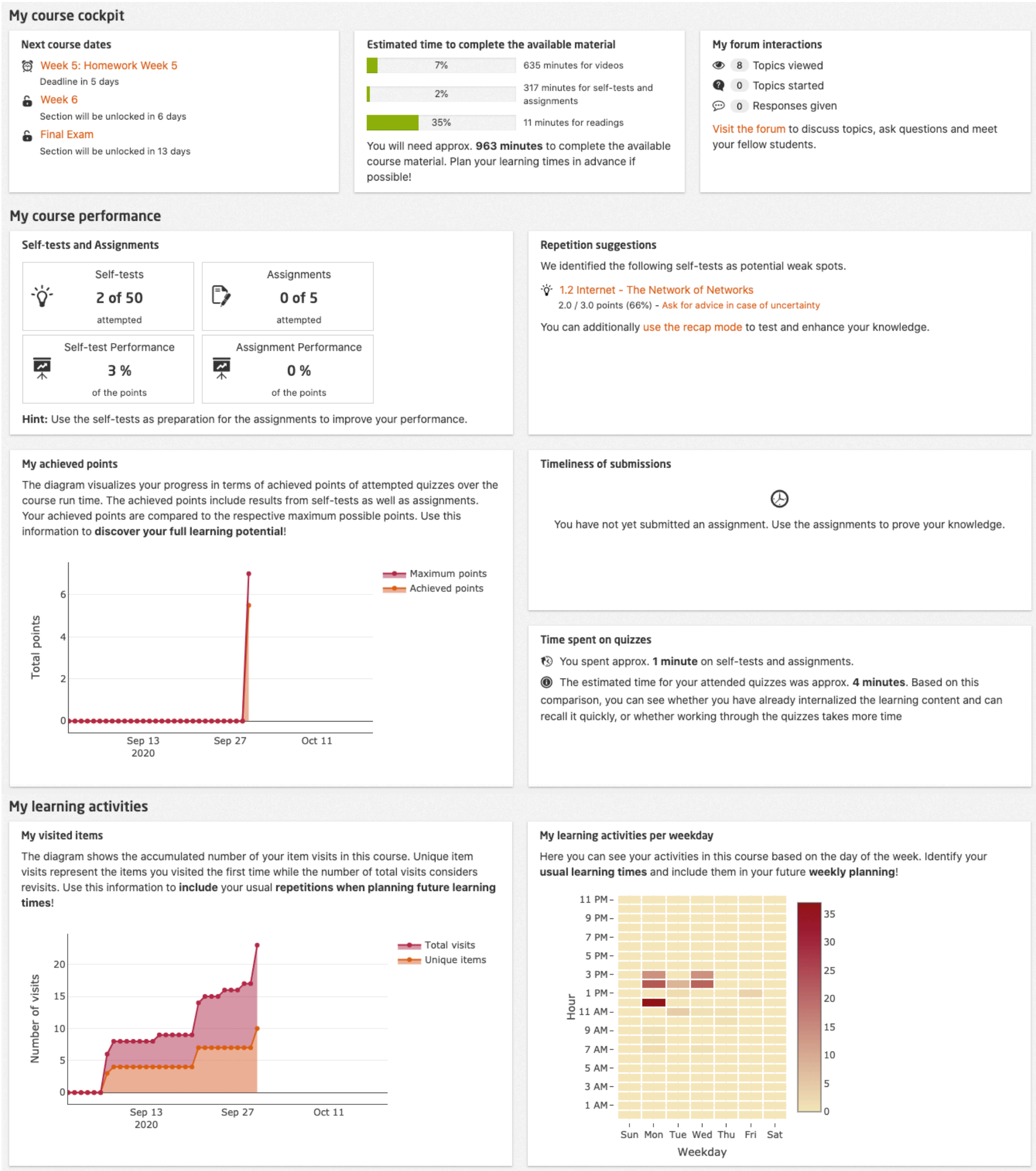


Figure 3.8: The Revised Learner Dashboard. The new Progress Overview is still displayed above it as shown in Figure 3.2, but is omitted in this figure since it was not further adapted for this experiment.

3.4.2 Evaluation

In the following, we evaluate research questions 2.3 and 2.4. For this purpose, we first present the examined courses and the methodology of the study. Then, we investigate the completion rates concerning the dashboard usage by performing an A/B/n test, and finally, we assess a survey to investigate whether the dashboard is supportive in the application of self-regulated learning.

Sample Courses

The experiment was conducted in five courses with different characteristics on openHPI, which we briefly present below. Further descriptive statistics for all these courses can also be found in Table 3.1. A Record of Achievement and a Confirmation of Participation could be achieved in all courses.

Table 3.1: Enrollments in Sample Courses for the Revised Learner Dashboard Experiment

	Enrollments			Shows			
Course	Start	Middle	End	Middle	End	Weeks	Language
internetworking2020	3,440	4,823	5,398	2,131	2,827	6	English
kieinstieg2020	9,268	10,515	11,284	5,492	6,856	4	German
digitalhealth2020	4,059	4,537	5,034	2,228	3,467	2	German
learningtheory2020	2,312	2,855	3,012	1,246	1,638	2	English
knowledgegraphs2020	4,812	6,081	6,511	2,994	3,527	6	English
Total	23,891	28,333	31,239	14,091	18,315	-	-

The first course, "A Half Century of Internet: How it works today" (internetworking2020[5]), was held in English and ran from September 1, 2020, to October 20, 2020. The course lasted six weeks and each week included a graded homework assignment. At the end of the course, participants could take a final exam. It covered technical aspects such as the structure, mechanisms, protocols, and applications of the Internet, but also its history and role in society. At the start of the course, 3,440 learners were enrolled.

The next course, "Artificial Intelligence and Machine Learning for Beginners" (kieinstieg2020[6]), was taught in German and scheduled for four weeks from September 8, 2020, to October 6, 2020. This course was also assessed by means of weekly graded homework assignments and a final exam. It was specifically—but not exclusively—aimed at young people and others interested without programming or technical experience. Concepts and processes of self-learning programs were taught, as well as an outlook on future developments, and ethical issues were discussed. 9,268 learners had enrolled at the beginning of the course.

The third course focused on the topic of "Digital Health for Beginners" (digitalhealth2020[7]). This two-week German-language course focused on basic concepts and the questions of why digital health is important for personal health and why digitization is a driver of innovation in

[5] https://open.hpi.de/courses/internetworking2020/

[6] https://open.hpi.de/courses/kieinstieg2020/

[7] https://open.hpi.de/courses/digitalhealth2020/

medicine. At the start of the course on September 22, 2020, a total of 4,059 learners were enrolled. It ended on October 6, 2020. For assessment, two graded tests were conducted at the end of the two weeks, but there was no final exam.

The fourth course, "Computational Learning Theory and Beyond" (learningtheory2020[8]), was aimed at students and experts with existing formal mathematical knowledge. The course discussed a model of binary classification in-depth and other learning models, as well as a broad overview of other approaches towards a theory of artificial intelligence. It lasted two weeks including graded homework and a final exam—and took place from October 6, 2020, to October 27, 2020. The English-language course had 2,312 learners enrolled at the start.

The last course on "Knowledge Graphs" (knowledgegraphs2020[9]) was held in English from October 27, 2020, to December 15, 2020. It spanned over six weeks with weekly graded homework and a final exam. Also, the course required a basic understanding of web technologies, mathematics, and databases to follow the taught topics about semantic technologies, knowledge representation, and symbolic artificial intelligence. At the start of the course, 4,812 learners were enrolled.

Methodology

The A/B/n test incorporated three different variants of the learner progress and dashboard. Therefore, the following test groups emerged. The round-robin scheduled group assignment was done at the first visit of a learning item in the respective courses.

Group 1: Learners assigned to this group were only able to use the old progress page (Figure 3.1). In the context of this test, the group served as a control group.

Group 2: This group got to see the new progress overview but not the additional learner dashboard widgets (Figure 3.2).

Group 3: This group got to see the complete revised learner dashboard, i. e., the new progress overview (Figure 3.2) and below the revised learner dashboard widgets (Figure 3.8).

At the end of the courses, we exported reports from the platform to compare the achieved points and visited learning items as percentage values between the groups. These two metrics are the main criteria for obtaining both certificates and are therefore the basis for calculating completion rates in the context of the HPI MOOC Platform. For the evaluation, only the 'shows at middle' are examined, i. e., the users who had a realistic chance of achieving an RoA. Also, users who were already part of a treatment group—i. e., group 2 or 3—in a previous course (chronologically ordered by start date) are excluded from the evaluation. Based on the percentage values of both metrics, we compare the three groups for statistically significant differences (p-value) utilizing a one-way ANOVA test.

Besides, we had invited group 3 to complete a voluntary survey after the courses were finished and the results were published. The complete response data is available in Appendix C.2. The survey is based on the Evaluation Framework for Learning Analytics (EFLA) by Scheffel [131], which provides eight quality indicators for LA applications grouped by the three dimensions data,

[8] https://open.hpi.de/courses/learningtheory2020/

[9] https://open.hpi.de/courses/knowledgegraphs2020/

awareness and reflection, and impact. The response options were defined using a Likert scale with numerical values from 1 (strongly disagree) to 10 (strongly agree). As the framework proposes, we calculate the average values for each question, then the dimensional scores by rounding the result of $((x-1)/9) * 100$ where x is the average value of a dimension, and finally the overall EFLA score as the mean of the three dimensional scores. The questions were phrased as follows:

DATA:
- D1: For the new Learner Dashboard it is clear what data is being collected.
- D2: For the new Learner Dashboard it is clear why the data is being collected.

AWARENESS AND REFLECTION:
- AR1: The new Learner Dashboard makes me aware of my current learning situation.
- AR2: The new Learner Dashboard makes me forecast my possible future learning situation given my (un)changed behavior.
- AR3: The new Learner Dashboard stimulates me to reflect on my past learning behavior.
- AR4: The new Learner Dashboard stimulates me to adapt my learning behavior if necessary.

IMPACT:
- I1: The new Learner Dashboard stimulates me to study more efficiently.
- I2: The new Learner Dashboard stimulates me to study more effectively.

Completion Rates

The size of the test groups varies from 183 to 1,524 users depending on the course, whereas the three group sizes within a course are almost equal. Across all courses, a total of 3,448 users are assessed in group 1, 3,440 users in group 2, and 3,433 users in group 3. Based on the average values in Table 3.2, it is visible that the courses had apparently widely varying levels of difficulty. Particularly, in terms of achieved points, the figures vary from 5.81% to 47.39%. The p-values of the ANOVA test between the three groups show no statistically significant differences in all five courses for both metrics—the visited items and achieved points. We also repeated the analysis for different user cohorts within the same test groups: (a) users for whom the corresponding course was the first taken on the platform, (b) users who indicated in their profile that their regular computer use was easy to intermediate, (c) users older than 50, and (d) users older than 60.

Table 3.2: Descriptive and Inferential Statistics for Completion Rates with Regard to the Use of the Three Dashboard Variants

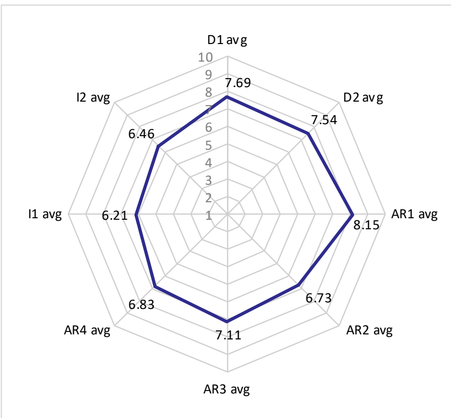
		Old Progress			New Progress			Complete Dashboard			ANOVA
Metric	Course	N	Mean	Std.Dev.	N	Mean	Std.Dev.	N	Mean	Std.Dev.	p-value
Visited Items (Percentage)	internetworking2020	572	42.35	41.32	580	40.95	40.74	564	43.27	41.25	0.631
	kieinstieg2020	1521	61.55	40.11	1524	62.25	39.20	1524	61.46	39.48	0.833
	digitalhealth2020	495	68.73	35.71	473	67.98	36.21	477	65.88	35.68	0.441
	learningtheory2020	183	30.00	29.85	201	29.58	30.23	188	30.27	29.56	0.974
	knowledgegraphs2020	677	34.16	37.95	662	32.82	36.65	680	32.47	36.46	0.676
Achieved Points (Percentage)	internetworking2020	572	30.81	39.24	580	29.56	38.69	564	30.97	39.21	0.798
	kieinstieg2020	1521	46.75	42.74	1524	46.38	42.23	1524	46.09	42.88	0.914
	digitalhealth2020	495	47.39	39.61	473	46.93	39.69	477	44.30	39.05	0.426
	learningtheory2020	183	6.61	20.24	201	5.83	17.94	188	5.81	17.09	0.893
	knowledgegraphs2020	677	15.06	27.18	662	13.63	25.64	680	13.79	25.92	0.549

We assumed that said user cohorts need more guidance in completing online courses. However, no statistically significant differences are discovered here either. Therefore, research question 2.3 can be answered unambiguously by stating that there are no differences in the completion rates of learners with regard to the use of the three dashboard variants. All three progress and dashboard variants are sufficient to achieve a certificate-based learning outcome. Of course, this is only one aspect of the overall learning experience and process. The previous study has already revealed the very well perceived usability and usefulness of the new learner dashboard. In the following, we explore how the dashboard supports self-regulated learning strategies, which is one of our main concerns.

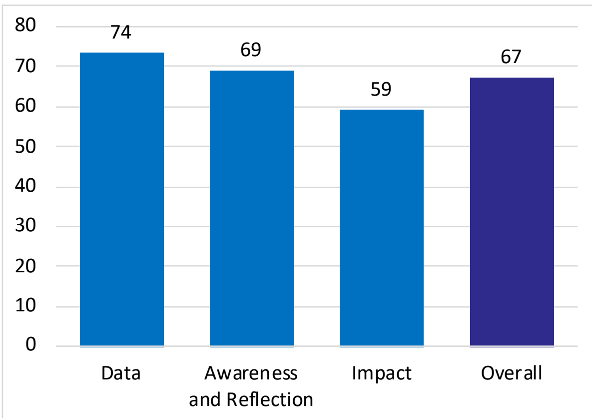
Supporting Self-Regulated Learning

The voluntary post-course survey received complete submissions from 296 learners of test group 3. The average numbers of the individual EFLA items in Figure 3.9a show that the dashboard provides the highest level of support in terms of awareness of one's current learning situation (AR1: 8.15). Users were also clear about what data (D1: 7.69) and why this data (D2: 7.54) is collected. The quality that the dashboard stimulates to reflect on one's past learning behavior (AR3: 7.11) was also rated positively. Moderately positive ratings were given to the dashboard's support in adapting the learning behavior if necessary (AR4: 6.83) and in forecasting one's own future learning situation (AR2: 6.73). Also rated moderately positively was the impact dimension, indicating that the dashboard stimulates to learn more effectively (I2: 6.46) and efficiently (6.21). In summary, all quality indicators were rated moderately positive to positive.

Our intention to support SRL with the learner dashboard (research question 2.4) is achieved especially for the strategy of self-evaluation by the stimulation of awareness (AR1) and reflection (AR3) of the learning situation and behavior. Also, strategic planning is partially encouraged by stimulating the adaptation of the learning behavior when necessary (AR4). The overall score of 67 of all EFLA dimensions in Figure 3.9b can be used to compare future versions of the dashboard with this one and to measure further improvements for the learner.



(a) Average Values of the EFLA Items



(b) Average Scores of the EFLA Dimensions

Figure 3.9: EFLA Results for the Revised Learner Dashboard

3.4.3 Conclusion

In this section, we implemented and evaluated another iteration of the learner dashboard. Thereby, the initial learner feedback and the results of the last study were incorporated. Hence, some learning analytics widgets on the dashboard were rearranged, removed, or provided with further explanations. After that, we first explored research question 2.3 of whether there are differences in the completion rates of learners with regard to the use of the three dashboard variants—the old progress page, the new progress overview, and the complete revised learner dashboard. For this purpose, an A/B/n test was conducted in five courses. The three variants were found to have no statistically significant differences in completion rates, and thus all are sufficient for certificate-based learning success.

However, one of our main concerns was on supporting self-regulated learning with the learner dashboard (research question 2.4). This was explored utilizing a survey that could be answered voluntarily by participants in the test group with the complete dashboard at the end of the courses. It was found that especially self-evaluation, by stimulating the awareness of the learning situation and reflection of the learning behavior, and partly strategic planning, by stimulating the adaptation of the learning behavior when necessary, are encouraged by the dashboard which matches our intentions. The next section examines how the SRL strategy of goal setting can additionally be enabled with technical means.

3.5 Exploring Learning Objectives in MOOCs

After having mainly addressed the self-regulated learning strategies of self-evaluation and strategic planning so far, this section focuses on goal setting to continue our efforts to provide technical support for SRL in MOOCs. Instead of measuring success based on certification and completion rates, researchers started to define success with alternative metrics recently, for example by evaluating the intention-behavior gap and goal achievement. Hence, we first examine how successfully learners achieve their self-reported learning objectives (research question 2.5). Afterward, a concept with a focus on technical feasibility and automation is outlined about how personalized learning objectives can be supported and implemented in MOOCs (research question 2.6).

3.5.1 The Status Quo

For the support of self-regulated learning in MOOCs certain approaches have been researched but only few related work is available which examines goal setting and achievement in MOOCs. This study aims to fill this gap. Therefore, this subsection investigates the current capabilities and limitations of goal setting and achievement on the HPI MOOC Platform before comparing the results with similar studies.

Sample Courses

To investigate the targeted and accomplished learning objectives of course participants, five courses are examined in this study (Table 3.3). These courses were conducted on openHPI. The taught topics are all based on the field of information technology and computer science. The required proficiency levels range from beginner to academic and professionals. In total, 25,801 learners were enrolled at course middle. A Record of Achievement was issued to those who have earned more than 50% of the maximum number of points for the sum of all graded assignments. A Confirmation of Participation was issued to those who have completed at least 50% of the course material.

Table 3.3: Enrollments in Sample Courses with Self-Reported Learning Objectives

	Enrollments			Shows			
Course	Start	Middle	End	Middle	End	Weeks	Language
javaeinstieg2017	7,127	9,242	10,402	6,610	8,015	4	German
javawork2017	3,881	4,112	4,336	1,482	2,096	2	German
searchengine2017	3,922	4,145	4,484	1,702	2,660	2	German
mainframes2017	2,635	3,026	3,396	1,670	2,115	6	German
imdb2017	4,683	5,276	5,825	2,402	3,128	6	English
Total	22,248	25,801	28,443	13,866	18,014	-	-

The first course, "Object-Oriented Programming in Java" (javaeinstieg2017[10]), was a four weeks course for beginners running from March 27, 2017, to May 14, 2017. Every week introduced different Java language features and object-oriented programming concepts with video lectures, followed by self tests and online programming exercises. Most of the programming exercises were graded for the final certificate. Additionally, an optional team peer assessment was conducted, where learners had the chance to gain bonus points. A total number of 9,242 enrollments was taken at course middle.

The next course was a two weeks workshop with the topic "Introduction into a Java IDE" (javawork2017[11]). This course was held from May 01, 2017, to May 15, 2017 and built upon the taught concepts of the javaeinstieg2017 course. Thus, a basic knowledge about the Java programming language was recommended. The first two weeks showed practical knowledge with lecture videos, followed by ungraded self tests. At the end, a graded peer assessment was conducted, which was the requirement to gain a certificate. 4,112 learners were enrolled at course middle.

The third course was a two week course as well, and addressed the question "How does a search engine work?" (searchengine2017[12]) from May 29, 2017, to June 20, 2017. The course was designed as an introduction to the topic for people outside the discipline, but also as a starting point for professionals and academic people, who want to get a first overview. The course structure followed the typical MOOC approach with consecutive videos and self tests. At the end, a graded exam was performed and 4,145 participants had been enrolled at course middle.

The fourth course about "Mainframes" (mainframes2017[13]) was held from June 05, 2017, until July 27, 2017. This six weeks course provided an in-depth perspective on mainframe architectures, application development, databases, security, and storage management. Thus, this course was mainly aimed at academic and professional people. Next to the video lectures and self tests, a

[10] `https://open.hpi.de/courses/javaeinstieg2017/`
[11] `https://open.hpi.de/courses/javawork2017/`
[12] `https://open.hpi.de/courses/searchengine2017/`
[13] `https://open.hpi.de/courses/mainframes2017/`

weekly graded assignment was conducted, as well as a graded exam at the end of the course. 3,026 learners were enrolled at course middle.

The "In-Memory Data Management" (imdb2017[14]) course dealt with the management of enterprise data in column-oriented in-memory databases and their inner mechanics. The course was running for six weeks from September 18, 2017, to November 18, 2017, and 5,276 learners enrolled in it. Due to the specific technical focus, the target groups were academics and professionals. This course was graded by a weekly assignment and a final exam.

In summary, the evaluated courses provide a well-balanced data basis with different course lengths, target groups, and proficiency levels, as well as different theoretical and practical examination modalities. All of them offered the two certificate types: an RoA and a CoP. Table 3.3 also displays the number of enrollments and 'shows'. Based on Hill's [52] definition of 'no-shows', i. e., learners who enrolled for a course but never viewed any content, an overall show rate of 53.78% at course middle was reached. Additionally, following the definitions of Renz et al. [107] a total completion rate of 29.02% and consumption rate of 52.30% are measured. When comparing the show rate and consumption rate, it can be seen that almost all active learners that enrolled before course middle visited more than 50% of all learning content and therefore gained a CoP.

Methodology

When accessing one of the courses for the first time, a welcome text with general information about the course was presented to the learner. The following item was an optional pre-course survey asking the learner about their primary goal for the enrollment into this course amongst other general questions. Based on the platform's feature set and available certificates, four mutually exclusive objectives were questioned:

Objective 1: I would like to receive a Record of Achievement in the end and learn the course content.

Objective 2: I am mainly interested in learning the course content. The Record of Achievement is not important to me.

Objective 3: I am only interested in selected learning units.

Objective 4: I just want to look around.

An overview of all criteria to achieve and to exceed the learning objectives is shown in Table 3.4. The achievement of objective 1 and 2 can be traced by course completion if a certain certificate was gained. To accomplish objective 1, an RoA must be reached. For objective 2 the assumption was made, that if a learner studied the majority of learning content (50%), a CoP was achieved. For the accomplishment of objective 3 and 4 a behavioral analysis based on user interaction events was conducted. To achieve objective 3, the user had to watch at least 1 video lecture. This is the base unit to measure if the user visited and interacted with any learning content since there is no platform feature available that enables the user to select the specific learning content they are interested in. For objective 4, the visit of at least 3 items was defined as the criteria to achieve the learning goal. This specific number was chosen because the first visited item was the welcome text when entering the course, the second was the survey itself, and the third item visit was the proof that at least one learning item was visited. These assumptions already show limitations of

[14] https://open.hpi.de/courses/imdb2017/

the platform regarding goal setting and evaluation. By following this approach, no additional self-reported data was necessary to determine goal achievement of all students that responded to the pre-course survey. All measurements are based on platform data, which reduces the influence of a survivorship bias. Therefore, it was not required for the evaluation that learners complete the course and participate in another post-course survey.

Table 3.4: Criteria for Learning Objective Achievement

Objective	Criteria to achieve objective	Criteria to exceed objective
Objective 1	Accomplish RoA	n/a
Objective 2	Accomplish CoP	Accomplish objective 1
Objective 3	Watch at least 1 video	Accomplish objective 1 or 2
Objective 4	Visit at least 3 items	Accomplish objective 1 or 2 or 3

Pre-Course Survey

The results of the pre-course survey of every course can be seen in Table 3.5. A total amount of 9,698 users provided their learning objective. In relation to the total number of 'shows at middle' (13,865) a response rate of 69.95% was reached. Between 22.52% and 36.03% stated that they want to receive an RoA (objective 1) with a total result of 26.63%. The majority of users (61.54%) were mainly interested in learning the course content without the need to gain an RoA and therefore chose objective 2 ranging from 54.41% to 65.80%. Between 3.62% and 5.41% selected objective 3 since they were only interested in selected learning units with a total result of 4.45%. Lastly, 7.37% stated that they only want to look around (objective 4) with a range from 5.94% to 10.74%.

Table 3.5: Pre-Course Survey: What is your primary goal for the enrollment into this course?

Course	Objective 1		Objective 2		Objective 3		Objective 4	
javaeinstieg2017	22.52%	1,006	65.80%	2,940	4.27%	191	7.41%	331
javawork2017	23.73%	342	64.33%	927	5.41%	78	6.52%	94
searchengine2017	32.18%	528	56.31%	924	4.75%	78	6.76%	111
mainframes2017	29.79%	319	55.18%	591	4.30%	46	10.74%	115
imdb2017	36.03%	388	54.41%	586	3.62%	39	5.94%	64
Total	26.63%	2,583	61.54%	5,968	4.45%	432	7.37%	715

Goal Achievement Analysis

When assessing the results of the pre-course survey, it is notable that only about one quarter of the users are interested in a graded performance appraisal and considerably more than half of the users are mainly interested in the content itself without the need of an RoA. This mirrors the varying learning objectives of lifelong learners since especially well-educated professionals form a large part of the learning community and not all of them are necessarily interested in gaining a certificate [34].

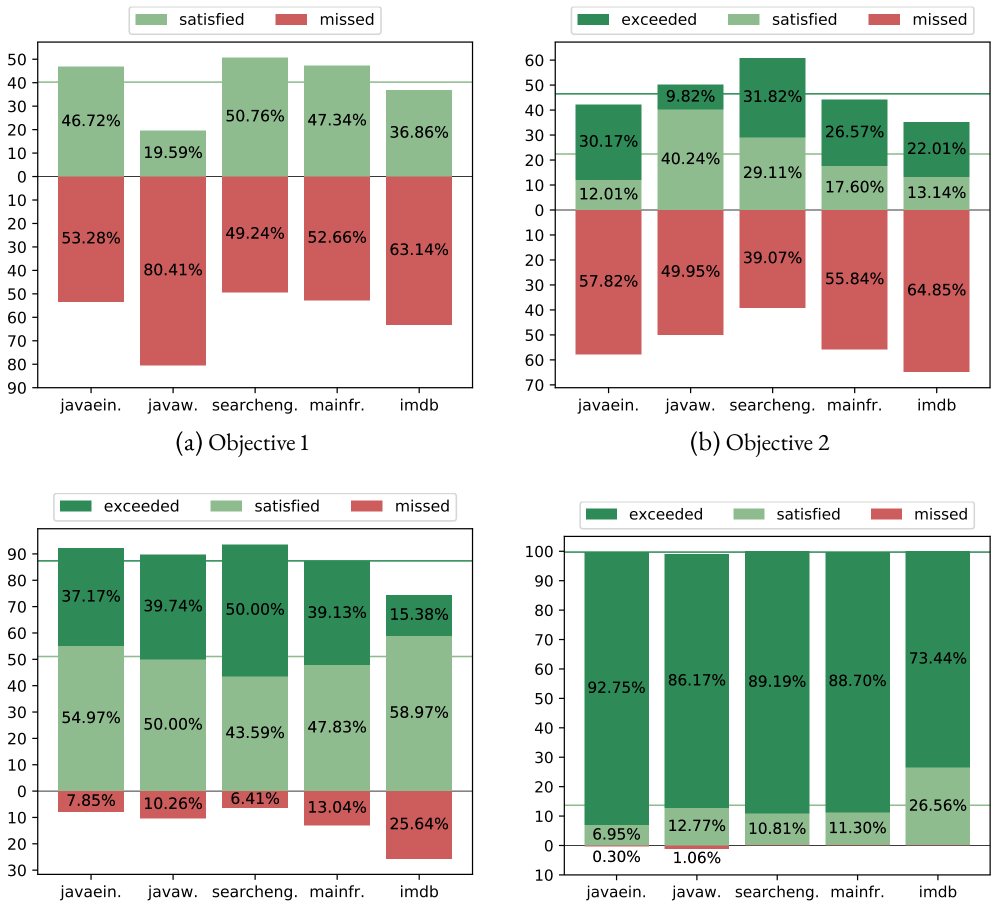
Few users stated that they are only interested in selected learning units or only want to look around. This may be related to the fact that at course start only the first week was available, and the remaining content followed week by week. This is a typical approach in MOOCs to foster discussions in the forum and support the mastery learning approach. Nevertheless, this shows the shortcoming that at the beginning of the course it is difficult to get an overview of all the content and topics that will be taught in the following weeks.

Table 3.6: Achieved Learning Objectives

Objective	Satisfied		Exceeded		Satisfied or Exce.		Missed	
Objective 1	42.55%	1,099	n/a		42.55%	1,099	57.45%	1,484
Objective 2	19.71%	1,176	26.11%	1,558	45.81%	2,734	54.19%	3,234
Objective 3	51.62%	223	38.19%	165	89.81%	388	10.19%	44
Objective 4	10.77%	77	88.95%	636	99.72%	713	0.28%	2
Total	25.09%	2,386	24.81%	2,359	49.90%	4,745	50.10%	4,764

In Table 3.6 the overall goal achievement is displayed. First, it can be seen that nearly half of the users achieved or exceeded their goals and the other half missed their objective. Also, the total satisfied and exceeded achievements are almost equally distributed. From this insight it can be derived that there is a large user group that either changes their goal during course runtime or drop out due to course related or non-course related barriers [49]. In both cases it shows the limitation that learning objectives cannot be set in a proper way which allows the user to also adjust them at a later point of time. Nevertheless, the results show a big range when comparing the different learning objectives with each other since the objectives with the highest achievement rate required much less course activity and vice versa.

Figure 3.10 displays the individual achievement rates for all courses, grouped by the defined objectives. These results are centered around a zero line in order to allow an easy comparison of the achieved learning objectives. Satisfying and exceeding a goal are stacked upwards, whereas missing a learning goal is stacked downwards. Additionally, the first horizontal line in the upper space marks the average mean of satisfying a goal, and the second line the average mean of satisfying and exceeding a goal combined. The specific mean values can be seen in Table 3.6. Compared with a standard deviation of 0.1155 for satisfying objective 1, it is notable that only the javawork2017 course shows a greater deviation. This can be attributed to the fact that this course was only graded by a peer assessment, which required much more effort than a typical multiple choice examination. The highest achievement rate was reached by the searchengine2017 course. This course was only graded by a single final exam without any weekly assignments, which reduced the required effort. The other three courses were graded by weekly assignments and a final exam. The achievement rates of objective 2 show a much higher variation, and objective 3 and 4 show overall high achievement rates, since these goals require less engagement. All in all, the individual achievement rates across the different courses point to the fact that goal achievement strongly depends on the course design, examination, and difficulty of different goals.



(a) Objective 1

(b) Objective 2

(c) Objective 3

(d) Objective 4

Figure 3.10: Achieved Learning Objectives per Course

Related Research

A sample size of five courses does not allow to draw general statements about goal achievement rates in MOOCs. Therefore, related and similar studies are presented in this subsection. A case study by Wilkowski et al. [168] about one course showed that 52.5% of their participants ($N = 20{,}977$) intended to complete their evaluated course with a (free of charge) certificate, from which 27% met or exceeded this goal at the end. The other learners preferred to learn new skills or explore the course content. Combined with these students who targeted smaller learning goals, a total number of 42.4% met or exceeded their goals at the end. The authors recommend to offer more personalized course designs based on students' goals, to move beyond the one-size-fits-all approach in MOOCs.

Another study with 53,128 enrollments across nine courses by John et al. [60] showed that only a few learners (0.64–1.24%) are interested in gaining a (charged) verified certificate to earn credits for their degree, on-the-job training, or job applications. From the participants who booked this certificate option, between 63.3% and 92.0% gained a certificate at the end, whereby the paid fee very likely increased their motivation. Henderikx et al. [50] examined the success of two MOOCs based on the intention-behavior gap. In the first course, 59% of their participants achieved or achieved more than initially intended ($N_1 = 65$). An even higher success rate of 70% was found in the second course ($N_2 = 101$). These results are based on a subset of learners who responded to the post-course survey leading to a survival bias. Nevertheless, they "underline the importance of individual perspectives" and recommend to consider that "individual goal achievement does not necessarily matches goal achievement from the institutional perspective." Other studies, which measured certificate achievement based on students' self-reported intention to complete a course, found completion rates between 22 and 29% [102, 170] or around 9% [71].

Discussion

To summarize the achievement rates regardless of the variation in the reported goal, a substantial percentage of students both meet or exceed, or miss their goals in MOOCs. The specific ratio is course-specific and likely depends on the course design and difficulty. Nevertheless, this and related studies show the importance to better support the presented strategies for self-regulated learning in MOOC environments. Thereby, different shortcomings are identified.

Currently, goal setting is mostly done with pre-course surveys. These can help the teaching team to get a broad insight into the overall motivation of their learning community. However, the learners have mostly neither the possibility to self-evaluate their learning process and outcome regarding their stated learning objective, nor be able to adjust their objective during the course runtime. Learner dashboards mostly focus on overall course completion [57], which does not reflect the objective of a large amount of learners, as the analysis has shown.

Also, the measurement of goal achievement is mostly done manually since the survey responses cannot be processed automatically. Sub-goals like the completion of a certain topic section or week are only provided if the teaching team prepares such survey answers. Generic answers like "I am only interested in selected learning units" as being used in this study include a certain bias since the learner is not aware of which selected learning units are available at all. Furthermore, also strategic planning has to be considered in a concept to better support self-regulated learning, next to goal setting and self-evaluation.

3.5.2 A Concept for Personalized Learning Objectives

This subsection outlines a comprehensive concept to further support goal setting, self-evaluation, and strategic planning in MOOCs. It builds upon of the previously identified capabilities and shortcomings of MOOC platforms in general but besides with a technical focus on feasibility and automation in the context of the HPI MOOC Platform. Henceforth, we use 'personalized learning objectives' as an umbrella term for the various aspects of this concept.

Goal Setting

Currently, goal setting is mostly done with pre-course surveys in many MOOC platforms. This has to be implemented as a course-independent platform feature, which offers the available learning objectives in a clear way. It needs to be studied if this has to be a mandatory step, e. g., as part of the course enrollment process, or an optional advice which can be shown to the user while browsing through the course. Therefore, a multivariate experiment can be used to examine if this is accepted and used by all learners or only by a sub-group. Also, it has to be possible to change the desired objective at any given time. By implementing such a feature, goal setting does not need to be maintained by the teaching team as a survey anymore. Also, this enables the evaluation of learning objectives within the platform itself to further monitor the learning progress based on them. In order to offer course-specific learning objectives, the learning content needs to be categorized and labeled first. This can be automated to some extent. Typically, knowledge transfer in MOOCs is based on video lectures and assessed with quizzes. Video segmentation is a well researched field, e. g., by visual transition detection [172] and can be further improved with outline extraction through analyzing the presentation slides [18]. Related quiz questions can be identified with natural language processing techniques. Also the course structure itself supports categorization, since it already offers an order and titles for each learning item and section. Of course, it is reasonable to do this manually first to test an initial prototype before implementing extensive automation.

The biggest challenge is a practical one: the availability of content. Course content is often provided and uploaded during the course runtime when users have already started learning. This is problematic with regard to the selection of learning objectives. It can be solved by either extending the scope of objectives as soon as new content is available and communicating this beforehand, or by supporting the teaching team to implement a complete structured course outline before course start without the actual content. A course builder tool enables to plan the weeks of a course ahead and helps to enrich them with goal metadata.

Self-Evaluation and Strategic Planning

The support of these two strategies has already been addressed partly with the learner dashboard (Section 3.3 and Section 3.4). Learner dashboards are a common practice to monitor learning progress and goal achievement. Therefore, the main focus is to make this tool compatible with learning objectives. In addition to the overall course progress, it has to be clearly visible which learning objective is being targeted and how far it has already been achieved. Furthermore, it has to be apparent which learning content belongs to the selected learning objective, but without depriving the learner of further content.

Also, strategic planning methods were identified as positive predictors of goal achievement [71]. Especially regarding learning objectives, technical support to plan time management and effort regulation come in handy. Therefore, an estimated time effort can also be provided for the learning objective and its contents. This has to be already visible during the goal setting, as well as during the self-evaluation on the dashboard to estimate the remaining effort. This is only a first approach to support strategic planning and more tools like custom reminders or planning prompts [4] can be provided in the future.

3.5.3 Conclusion

This section introduced the potential of personalized learning objectives in MOOCs to further support self-regulated learning. PLOs shift the focus from completion-centered success rates based on gained certificates to individual course goals which better accomplish the needs of lifelong learners. Therefore, the current status quo of learning objectives in MOOCs was examined with an observational study of five courses regarding how well learners in MOOCs achieved their initially self-reported learning objectives (research question 2.5). The results and the comparison with similar studies showed that goal achievement rates are course-specific and likely depend on course design, examination modalities, and difficulty. In total, almost 70% of all active learners at course middle provided a course objective ($N = 13{,}865$). 49,90% of learners achieved or exceeded their goals, but also the effort required for a specific goal strongly affected the achievement rates. Nevertheless, technical support for goal setting and achievement is rare. Most studies rely on self-reported data from user surveys, which does not allow to provide feedback based on the selected goals and also the teaching team cannot draw any further conclusions about progress and success afterward.

From a pedagogical perspective, self-regulated learning was identified as a key skill set for learner achievement, especially in online learning environments with low guidance and support such as MOOCs. Therefore, the strategies goal setting, self-evaluation, and strategic planning were outlined with possible implementations in a concept to support personalized learning objectives in MOOCs (research question 2.6). Here, the focus was set on technical feasibility and automation to provide such functionality on a platform level instead of individual course designs by different teaching teams. This paves the way for further research in this field and supports the transition from a one-size-fits-all approach in online learning at scale to a more individual learning experience tailored for the needs of lifelong learners. The implementation and examination of this concept is addressed in the next sections.

3.6 The Integration of Personalized Learning Objectives

The concept for personalized learning objectives in MOOCs, which was outlined in the previous section, is now technically integrated into the HPI MOOC Platform and evaluated (research question 2.7). Furthermore, it is assessed with a mixed-method approach with regards to the perceived acceptance and usefulness (research question 2.8). The learners' acceptance is examined with an A/B/n test in two courses. Also, a survey is conducted to gather further feedback about the perceived usefulness, next to the acceptance.

3.6.1 Foundations on Learning Objectives

In this study, we continue our efforts to better support SRL in MOOCs. In doing so, we strive for a maximum degree of automation to ensure scalability and, ideally, to avoid additional effort for course instructors. This section explains the definition of learning objectives to achieve a common understanding of these terms and emphasize their benefits for learners.

Learning Goals and Learning Objectives

The terms learning goals and learning objectives are often used interchangeably as both describe the intended outcome of a learning process. However, the following distinction can be made [99, 149]. A learning goal is a broad statement of what a learner is be able to do at a certain time. It provides an overview describing a rather wide range of knowledge and skills a student acquires and is therefore usually not explicitly measurable. In contrast, learning objectives have a narrow focus, describing specific and discrete units of knowledge and skills being acquired. These objectives are the results of short time activities that can be achieved by following a certain number of steps. Consequently, they are specific enough to be observable and measurable.

In pedagogy, learning objectives are typically classified and created using models like Bloom's (Revised) Taxonomy [74]. Another well-known approach to define objectives is the *SMART* acronym [29]—objectives have to be specific, measurable, achievable, relevant, and time-bound. A learning goal thus can comprise multiple learning objectives.

Definition of Learning Objectives

Learning objectives describe the desired outcomes of learning processes. A learning outcome can be the acquisition of subject-specific knowledge but also the development of skills, abilities, and competences. This distinction is important since formally learned knowledge does not necessarily enable learners to adequately apply the knowledge in a specific situation. Proper assessment methods have to be in place to be able to measure and verify the attainment of objectives. Only if the outcome is measurable, a quantified decision about the level of success can be made and provided as feedback to the learner, which is desired to enable self-regulation.

The predominant xMOOC concept focuses on the acquisition of subject-specific knowledge, mainly imparted with pre-recorded video lectures. Courses group the content by specific topics and typically address different smaller thematic units. In contrast to the predominant orientation towards the completion of the course, individual objectives can be understood as completing certain parts of the course material by offering a form of optional personalized pathways. Therefore, we define the completion of these thematic units, built upon the single learning resources as the basis for learning objectives since they represent the smallest unit of imparted knowledge within a course. This view is compatible with the xMOOC concept and reflects the needs of lifelong learners, who are primarily interested in gaining specific knowledge [89]. The verification of the acquired knowledge is possible through the provided exercises. Furthermore, personalization is achieved by offering different didactically appropriate objectives per course, created by the teaching team and course instructors, from which the learner can select one if desired and follow it individually.

3.6.2 Related Work

So far, SRL has been extensively researched in formal classroom settings and also in traditional online learning. The results show that SRL is an important factor of successful learning [13]. Over the last years, it has increasingly gained attention in the context of learning in MOOCs. A common focus in literature is on identifying how learners apply SRL strategies and which strategies are most effective with regard to the learner's behavior and learning outcomes as this forms the basis for a proper (technology-based) support of these strategies [71, 76]. This support is crucial since learners differ in their abilities and motivation to regulate their learning [88, 95]. Different authors proposed design guidelines and patterns to facilitate SRL in MOOCs [79, 97].

Despite the recognized importance of goal-orientation in MOOCs, goal setting has been realized on the basis of pre-course surveys. For example, Wilkowski et al. [168] used a survey to enable learners to set the initial goal, and a post-course survey in combination with clickstream analysis to evaluate goal attainment. Also utilizing questionnaires, Henderikx et al. [50] analyzed goal achievement based on the intention-behavior gap. In the previous study, we also examined the intentions of learners in five courses using a survey for setting objectives and utilizing LA capabilities to evaluate their achievement. All of these studies show that a certain number of learners achieve their (initial) learning objectives while there is also a specific portion of learners which exceed or underachieve their objective. The actual achievement rates depend on the specific courses in terms of their design and difficulty as well as the required effort to complete the individual objectives.

Since the current capabilities in terms of goal setting are not sufficient to actively support learners, we have identified the following key requirements for future work: First, MOOC platforms have to offer the possibility to set an objective within the platform itself so that learners can self-evaluate their progress towards the achievement of their objective. This enables the automatic calculation of the achievement and allows a more fine-grained definition of objectives. Second, learners have to be able to adjust the objective as the course progresses.

Interventions for strategic planning are rare, too. For example, a time planner was integrated into a MOOC platform enabling students to schedule their next study sessions [4]. Asking learners to describe how they plan to study for the upcoming week, supplemented by another prompt at the end of the week instructing learners to reflect on the success of their plan, did not produce significant improvements [24]. In contrast, an intervention of Yeomans et al. [170], who also provided a planning prompt at course start, positively influenced the learners' course completion.

3.6.3 Concept and Implementation

This subsection describes the integration of PLOs into the HPI MOOC Platform as a proof of concept for blurring the completion-oriented structure of current MOOCs (research question 2.7). It introduces support for the SRL strategies goal setting and selected aspects of strategic planning. To overcome the current limitations regarding goal setting in MOOCs, the fundamental idea is to provide a tool, integrated into the platform, that empowers learners to actively choose and set a specific learning objective.

Course-Level Learning Objectives

Based on the presented definition of learning objectives, three types of objectives emerge that have to be supported on course-level. These allow different levels of engagement and are in line with learners' objectives reported in literature.

Striving for course completion to receive a certificate is still the intention of many learners participating in online courses, and therefore this represents the first type of objectives. In terms of the definition of learning objectives above, the completion of a course can be seen as the completion of a broader topic unit that comprises all course material. In the HPI MOOC Platform, course completion is rewarded with a so-called Record of Achievement. In addition to this, a Confirmation of Participation can be received when a specific proportion of the learning material is visited. Both types of certificates are possible objectives to be considered. Although these are not objectives according to pedagogical theory in a narrow sense, this type of objectives is reasonable as a simplification approach in the MOOC context.

In addition, different thematic units can be derived and offered as learning objectives. Beyond the thematic focus, another option provided by these smaller objectives is to adapt to proficiency or time aspects. While some objectives can go into detailed aspects of the courses, others can only give an overview to accommodate learners with limited time or missing prior knowledge of a topic.

Last, since there is a large number of learners just having a look at the course to find out whether it suits their needs and if it is worth pursuing, an objective approaching the course exploration may be desirable for specific courses. Course exploration is a typical pattern that has been identified by different authors [65]. Similar to the first type, this is not a pedagogical learning objective in the narrow sense but enables the teaching team to track this intention.

All conditions of the definition of learning objectives are fulfilled for all three types, which is why we use this term in the scope of this work. However, the definition of learning goals also applies to the first and second type, as the definitions of objectives and goals are not mutually exclusive. In addition, the teaching team determines the granularity and number of learning items that belong to the third objective type.

Learning Objectives Model

As presented in the related work, goal setting so far has been realized with surveys, which has several weaknesses. The goal of this work is to implement learning objectives as a platform feature to be able to determine the learners' achievement of objectives automatically. Therefore, a model for learning objectives needs to be defined. For the model, two requirements are particularly important. Although our approach is initially limited to course-level learning objectives, it has to be easily extensible to platform-wide objectives. Further, the model needs to be flexible in terms of creating learning objectives for various courses that differ in their structure, the type of included learning material, and the actual content. With the definition of learning objectives above, learning items, mainly videos and quizzes, form the basis for learning objectives. In courses, items are grouped by sections, i. e., by the weeks of a course. When considering platform-wide objectives, learning objectives are likely to be defined across courses and thus also contain items or sections from different courses. Besides, learning objectives can be organized hierarchically, e. g., to represent sub-steps and aspects for the mastery of a larger objective. Figure 3.11 shows a simplified version of this logical composition of objectives.

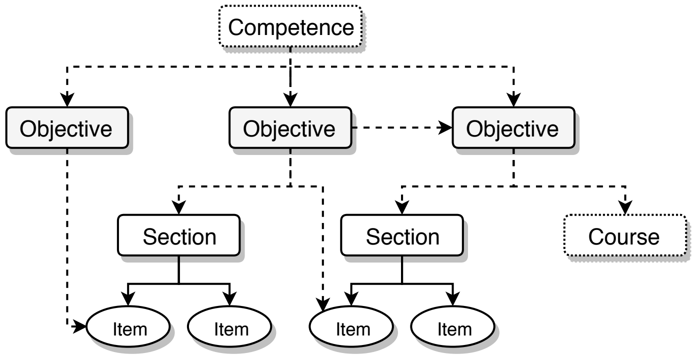


Figure 3.11: The Logical Composition of Learning Objectives. Bold arrows represent course structure dependencies, while dashed arrows indicate the logical relation between the concepts.

A natural consideration for creating learning objectives is to reuse the structural elements, i. e., course sections. However, sections are designed as a part of a course and thus are defined on a different granularity level than objectives and are likely not to fit an objective's focus and intention. Consequently, learning objectives need to be built on the learning items as the most fine-granular structures within a course. As a second issue, several items often belong to one topic area or are useful if studied after each other and can thus be grouped on a logical level. Therefore, it is necessary to introduce an abstraction layer, the so-called learning units, to enable both grouping of content and automatic calculation of goal achievement. Learning units can encompass several items and classify them into knowledge acquisition and knowledge examination items. An objective can then be created from different learning units. This results in very high flexibility accomodating many use cases. To enable platform-wide objectives, objectives need to be defined to be valid in a specific context rather than bound to a course directly. Figure 3.12 shows the resulting conceptual model for learning objectives.

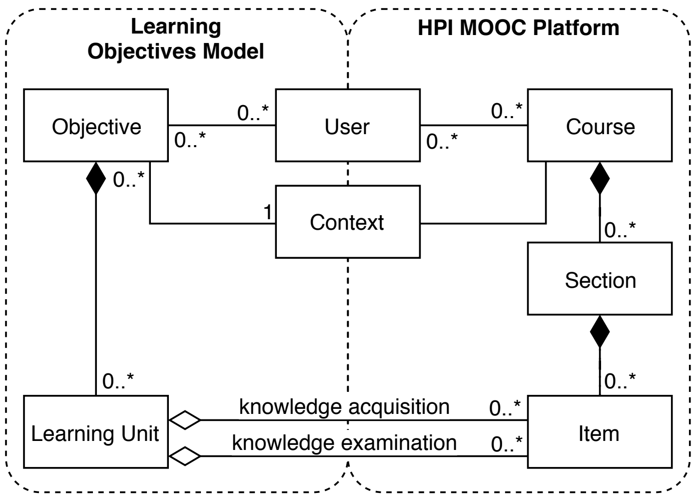


Figure 3.12: The Learning Objectives Model and the Relation to the HPI MOOC Platform

Integration into the Platform Architecture

In the current system based on a service-oriented architecture, the course service contains the domain logic concerning the courses and their item structure [158]. Consequently, the learning objectives can be implemented in this service as they depend on the course item representation and rely on its data. However, some arguments contradict an extension of the course service. The learning objectives extend the course domain but do not represent core functionality being critical for the platform. Besides, decoupling the new functionality from the course makes it easier to extend the concept with platform-wide objectives or competency models in the future. For these reasons, the learning objectives are integrated by creating a new learning objectives service. Its core purpose is to store and manage the learning objectives and calculate their progress and completion. Figure 3.13 shows the concept of the new service.

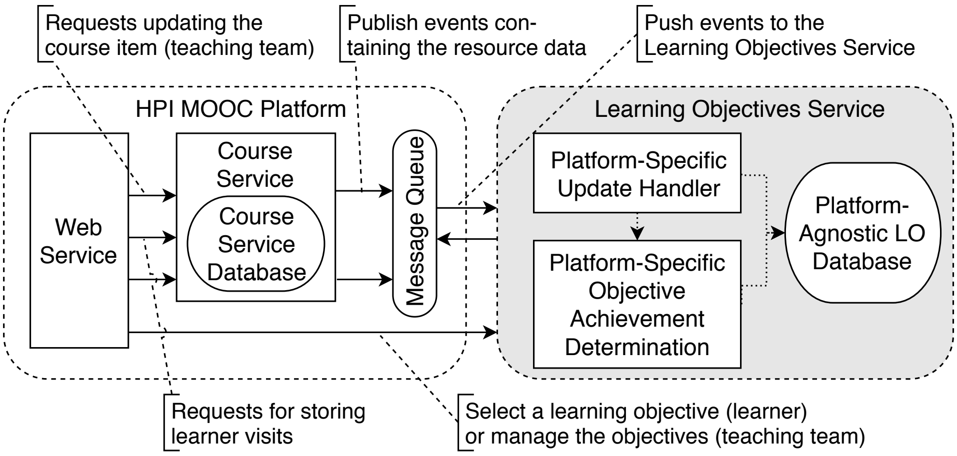


Figure 3.13: The Concept of the Learning Objectives Service

Selecting Objectives

A user interface (UI) must be created to empower learners to decide on a learning objective for a course. In the following, we present our concept for providing learners with a list of available learning objectives and related information enabling informed decisions. Three aspects decisively influence the selection of an adequate presentation concept. A first decision that has to be taken is whether the selection of a learning objective is optional or compulsory. While compulsory objectives may force learners to reflect on their intentions and thus provide an opportunity to improve learning, such an approach restricts the open nature of learning in MOOCs and is likely to upset learners as they may want to stick to the learning path and platform features they are used to. Personal learning objectives as introduced in this work are an extension of the traditional learning. As such, they have to be explorable as an optional feature.

In addition, it needs to be determined when exactly learners are allowed to choose a learning objective. New courses are usually announced weeks before their start, which leads to many learners enrolling impromptu but then never showing up for a course. Further peeks of enrollment are reached when actively advertising a course via email or social media. Consequently, an interesting option is to allow learners to select their learning objective directly after enrolling for a course. Goal setting at this point may increase engagement and help to build a stronger relation with a subject. A second possibility to select an objective is directly at course start when first visiting the learning content. This can help learners to focus on a specific part of the course and to not spend time on personally less relevant content. Last, other options include the selection of an objective after the learner worked on some items, e. g., three items so that the user has already completed the introduction and typically the first video and quiz.

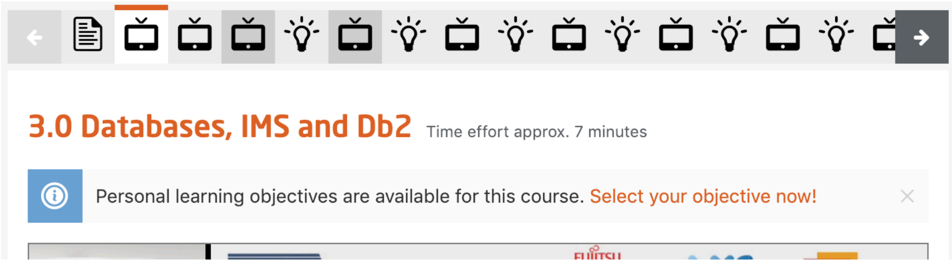


Figure 3.14: The Infobox Indicating the Availability of Learning Objectives

A third decision must be made on how to present the objective selection as this may also affect the adoption of learning objectives as well. Thus, multiple options are implemented. First, an infobox indicating that learning objectives are available for the course is provided. This infobox, as depicted in Figure 3.14, is added to the top of the learning item pages and thus prominently visible for the learner when working on the course material. Nevertheless, it is optional and the user has to click on a link to open the objective selection. Moreover, a modal is used for automatically prompting learners with the objective selection. It is explained in detail in the next subsection. For our purpose, the automatic display is an adequate approach to attract the students' attention, which is relevant since the feature is entirely new to the platform. To avoid learner frustration, the modal can be dismissed. Afterward, it does not show up again for the course automatically. The same applies to the infobox. The corresponding experiment setup and the results are detailed in Subsection 3.6.4.

Objective Selection Modal

As a basis for the objective selection, a modal[15] is used for two reasons. With the modal, learners do not have to leave the current page and thus remain in the context in which they can continue after selecting or refusing an objective. Another advantage of the widget-like presentation is the reusability for different pages and use cases, e. g., to allow the learner to change the objective while working on the course material.

The modal itself contains information on how the selection of a learning objective affects the learning process and lists the different available objectives as shown in Figure 3.15. The objective details, which are necessary to enable learners to make an informed decision, can be expanded by clicking on an objective. Since time is an important factor, the students are presented with information about the time needed to complete a particular objective. Besides the total time effort for

[15] A graphical overlay window also called dialog or pop-up.

an objective, the distribution of learning material in terms of the estimated time to study material constituting the learning objective is depicted aggregated by its type. With the information provided, the objectives can be compared in terms of the topics covered, but also in terms of the approximate effort to be invested for the successful completion of the objective. The stated type of material can help learners to select an appropriate objective for the educational background or proficiency level, e. g., to focus on videos rather than programming tasks or vice versa. In this early phase of introducing learning objectives in MOOCs with the purpose of evaluating their acceptance, the selection is currently limited to one objective per course at the time. After selecting a learning objective, the user is prompted with a confirmation of the selection as shown in Figure 3.16. The user can then choose to be directed to the first learning item part of the objective to immediately start learning.

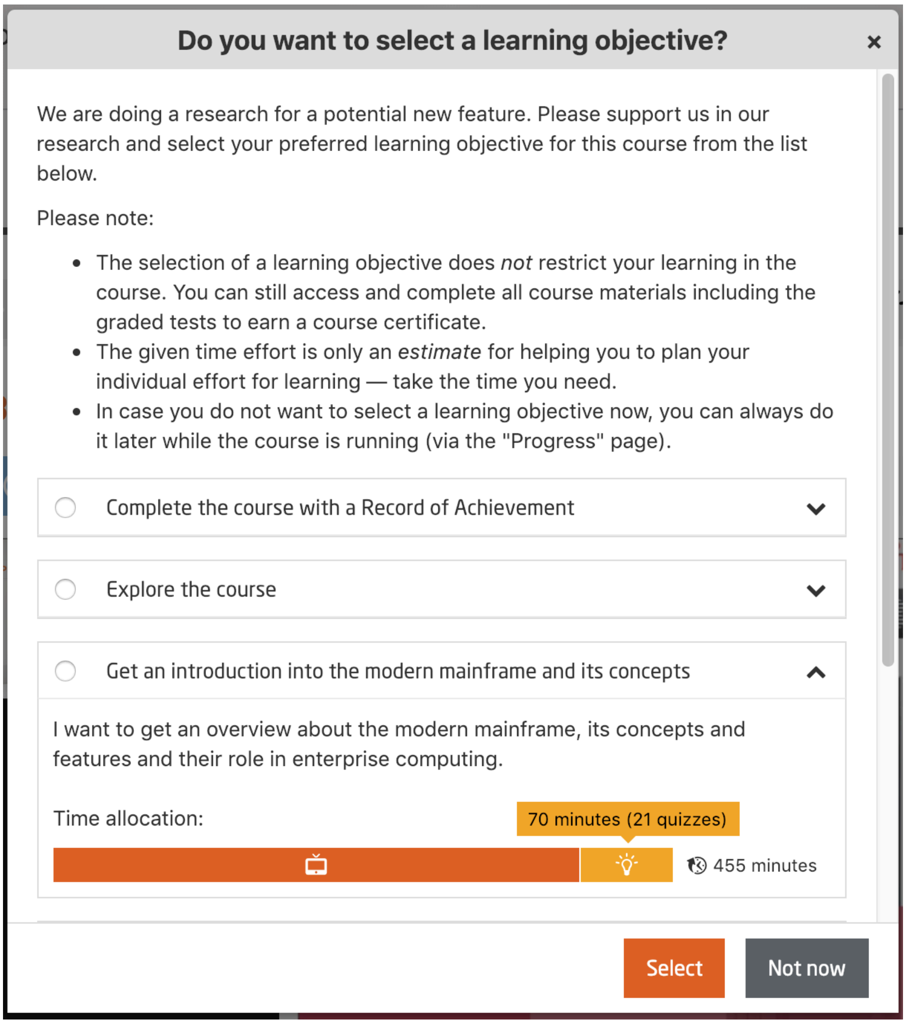


Figure 3.15: The Objective Selection Modal. It lists all available objectives and the respective details.

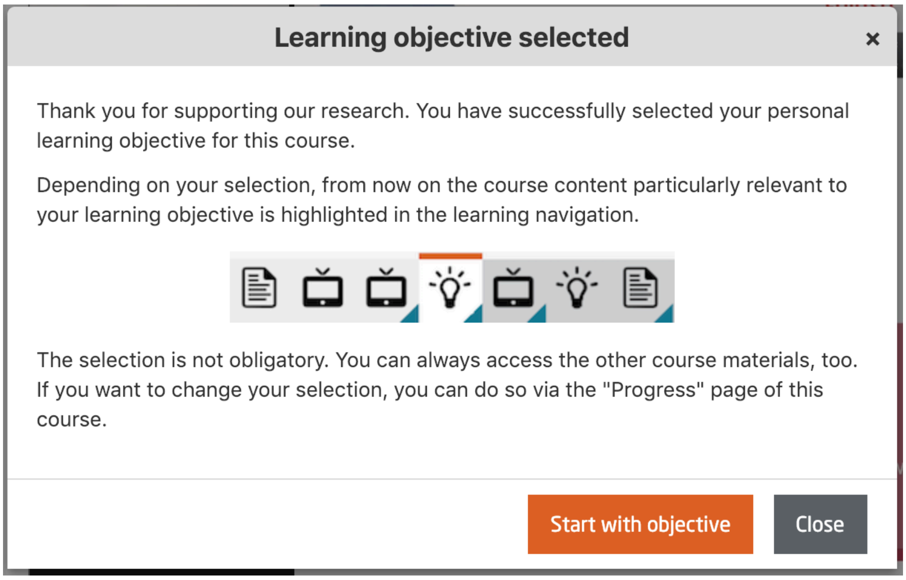


Figure 3.16: The Objective Confirmation Modal. It is displayed after selecting a learning objective.

Learning Objective Adaption

A requirement emerging from the deficiencies of goal setting with surveys is that learners have to be able to change their selected objective at any given time during the learning process in the event of changing personal conditions. For example, learners may face time constraints or become more or less interested and engaged and hence want to change their objective. Therefore, the personal learning objective can be reviewed and changed on the progress page of a course as shown in Figure 3.17.

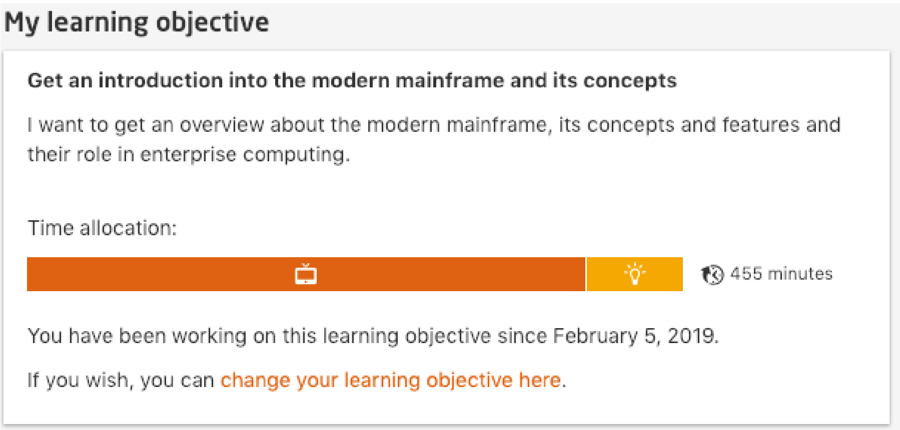


Figure 3.17: The Learning Objective Adaption. An excerpt of the progress page of a course where the selected learning objective can be changed.

Learning Process Guidance

With the learning objectives, the traditional course structure can be opened up. To allow students to quickly identify relevant content, the respective learning resources according to one's selected learning objective are highlighted. Specifically, the course items are indicated with blue triangles in the course navigation, which can be seen in Figure 3.18. Additionally, a textual clue is given. Thereby, learners can see where to start with the objective and the material to particularly focus on. Since there are cases where an objective does not start in the first week of a course or skips a week and thus no learning items can be suggested by then, the sections containing relevant learning material are emphasized as well. Beyond this indication of relevant content, the time effort information is provided in the course navigation to help students priotizing their learning activities and scheduling learning sessions accordingly, i. e., it is provided to support strategic planning.

As can be seen, the selected approach of guidance does not pull learners entirely out of the classic course setting. Instead, the course structure is maintained, i. e., the learning content is not rearranged and learners can decide whether they follow the defined course structure or adapt to a more flexible learning path according to the selected objective. This is important to motivate learners to do more as initially intended since they can also view and access the course material not being part of the objective. The decision on a learning objective consequently does not restrict the learning content in any way and the control of learning is still with the learner. In the future, this guidance can be explored further and extended with additional suggestions.

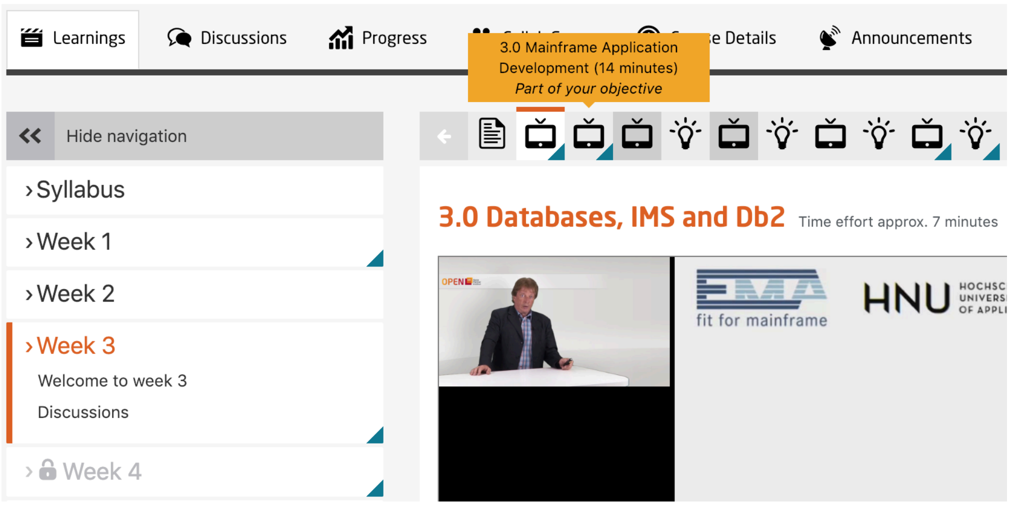


Figure 3.18: The Learning Process Guidance. It highlights the relevant content for a selected objective.

3.6.4 Evaluation

This subsection presents the evaluation of the integration of PLOs into a MOOC platform. Therefore, the A/B test framework of the HPI MOOC Platform [104] is used to examine the acceptance of learning objectives. Since the usability and perceived usefulness of a user-facing tool are critical factors for its adoption, additionally a survey is carried out to quantify these factors.

Sample Courses

The study was conducted in two courses—one each on openHPI and openSAP—to explore the results for different learner demographics and backgrounds.

On openHPI, the course "Mainframe - Crucial Role in Modern Enterprise Computing" (mainframes2018[16]) was selected, which started on November 5, 2018, running for six weeks until December 17, 2018. It covered different aspects of mainframes including its concepts and features like mainframe architecture, operating systems, application development, and also provided industry examples. 2,270 learners were enrolled at course start and the course language was English. The following objectives for the course were derived:

1. *Complete the course with a Record of Achievement*. This objective comprised all course material including the weekly assignments and the final exam.

2. *Explore the course*. Learners who did not know whether or not the course is interesting to them could choose this objective to take a look at the course.

3. *Introduction into the modern mainframe and its concepts*. The introduction covered the main parts of the first week and further mainframe concepts to give an overview of the topic.

4. – 7. Four more objectives covered the introductory content but then focused on specific aspects of the course: mainframe architecture and hardware, application development, database and transaction processing, and examples and scenarios from industry.

On openSAP, the course "Intelligent ERP with SAP S/4HANA Cloud" (s4h12[17]) was chosen which ran for four weeks from November 7, 2018, until December 6, 2018. The course presented SAP's intelligent cloud ERP solution for SAP S/4HANA Cloud and showcased use cases for different application areas. 13,512 learners were enrolled at course start and the course language was English. The objectives for the s4h12 course were defined by the openSAP teaching team as follows:

1. *Complete the course with a course certificate*. This objective aimed learners who want to complete the entire course including all weekly assignments and the final exam.

2. *Focus on the introduction into intelligent ERP*. The objective included the first week's material providing the technical foundation and selected videos of subsequent weeks, e. g., giving a general outlook for the course topic.

3. – 7. Five additional objectives covered the learning material of the introduction objective as well as the respective material detailing a specific use case: finance, procurement, project management, sales, and manufacturing.

Acceptance of Learning Objectives

In the following, the design and evaluation of the performed A/B/n test are explained to ascertain the acceptance of learning objectives and study research question 2.8. Specifically, the implemented presentation alternatives for the objective selection are examined to determine their effective-

[16] `https://open.hpi.de/courses/mainframes2018/`

[17] `https://open.sap.com/courses/s4h12/`

ness in regard to engage learners to set an objective. In addition, the focus is on the analysis of the learners' preferred choice of objectives. This also helps to give a comprehensive answer to research question 2.7 by identifying the best integration option for the new feature.

Experiment Setup The A/B/n test incorporated three different realizations of the objective selection. This resulted in four test groups, which were assigned at enrollment and formed on a round-robin schedule.

Group 1: Learners assigned to this group were not able to select a learning objective. In the context of this test, the group served as an independent control group.

Group 2: This group got to see the objective selection modal directly after enrolling in a course.

Group 3: Learners of this group were automatically prompted with the objective selection modal when visiting the learning content the first time after the course had started. Since a first decision can be to dismiss the modal, the infobox was added at the top of each item page, to reopen the modal with a click on it.

Group 4: In contrast to the other groups, the learners assigned to this group only saw the infobox at the top of each item page and therefore explicitly had to click on the link in the infobox to see the objective selection modal.

After the selection of an objective, the corresponding learning content was highlighted in the course navigation. Learners could review or change their objective on the course progress page.

Analysis and Discussion The central questions to answer are whether learners do select objectives, which tested selection alternative is best suited for the examined platforms, and which type of objectives is preferred. Table 3.7 shows the proportion of learners who selected learning objectives for the examined openSAP and openHPI courses respectively. Since only members of the groups 2, 3, and 4 were allowed to choose an objective, the results for these groups are displayed. Learners assigned to group 2 were prompted to select an objective immediately after the enrollment for the course. Because the number of additional enrollments after the start of a course

Table 3.7: Descriptive Statistics of the Learners With and Without Selected Objectives

			With Objective		Without Objective	
Course	Group	N	Learners	Quota	Learners	Quota
s4h12	2	1,010	537	0.532	473	0.468
	3	2,077	1,027	0.494	1,050	0.506
	4	2,074	398	0.192	1,676	0.808
	Total	5,161	1,962	0.380	3,199	0.620
mainframes2018	2	116	68	0.586	48	0.414
	3	322	189	0.587	133	0.413
	4	323	106	0.328	217	0.672
	Total	761	363	0.477	398	0.523

Note: control group 1 contained 2,001 (s4h12) resp. 292 (mainframes2018) learners.

is limited, this group contains fewer learners than the other groups, which also include learners who enrolled prior to the A/B/n test start and then showed up during course run time. In general, a considerable portion of 38% respectively 47.7% of the learners sets a personal learning objective for a course demonstrating the interest of learners to select a personal objective and confirming findings of the related work. However, there are differences between the groups, i. e., the presentation alternatives for the selection, which are consistent across both courses. In the groups where the selection modal was shown, about 49.4% up to 58.7% of the learners selected an objective while the more subtle alternative of showing an infobox attracted noticeably fewer learners. An analysis of demographic variables, i. e., the learners' age and gender, did not yield relevant differences.

Table 3.8: Descriptive Statistics of the Alternative Used by Learners for Selecting Their Initial Objective

			Modal		Info Box		Progress	
Course	Group	N	N_M	Quota	N_I	Quota	N_P	Quota
s4h12	2	537	529	0.985	-	-	8	0.015
	3	1,027	885	0.862	123	0.120	19	0.019
	4	398	-	-	356	0.894	42	0.106
mainframes2018	2	68	64	0.941	-	-	4	0.059
	3	189	155	0.820	30	0.159	4	0.021
	4	106	-	-	95	0.896	11	0.104

To further examine the effectiveness of the selection alternatives, for each group the different variants are compared regarding the learners' initial selection of an objective. The results, which are shown in Table 3.8, emphasize the importance of the objective selection via the modal as the majority of learners assigned to groups 2 and 3 decided on an objective when prompted with the modal. For the third group, the infobox additionally served as an important second step to attract many students who first dismissed the modal but then decided to set an objective. For both the second and third group, only a small portion of the learners had set the objective via the progress page. This suggests that it makes sense to explicitly encourage learners to use the feature instead of relying on its discovery by learners. Besides, there is relevance for offering different places and opportunities for selecting the objective. In sum, the third option of showing both a modal and an infobox is best suited to nudge learners to set a personal learning objective. Although the option of prompting learners when enrolling for a course seems promising as well, the current course creation process does not allow to largely apply this selection alternative.

A relevant question concerning the selected objectives is whether the learners prefer to complete the entire course or rather choose a specific topic unit of interest. Table 3.9 compares the respective results for both courses revealing similar overall tendencies. For simplicity reasons, the results for the objectives focusing on particular topic units of a course are aggregated. In both courses, the majority of learners intended to complete the course (65.7% and 55.9%) while about thirty percent of the learners either wanted to get an overview about the course or focus on a more specific aspect of the course. In contrast to the mainframes2018 course, where the shares are evenly distributed among these two groups, the s4h12 learners tended to choose the latter. These findings confirm that the interests and intentions for a course vary remarkably and learners do not solely

Table 3.9: Descriptive Statistics for the Distribution of Selected Objectives. The objectives covering individual topics are aggregated.

	s4h12		mainframes2018	
Objective Type	Learners	Quota	Learners	Quota
Completion	1,290	0.657	203	0.559
Topic	437	0.223	51	0.140
Introduction	235	0.120	51	0.140
Exploration	-	-	58	0.160
Total	1,962	1.000	363	1.000

focus on course completion but also prefer individual learning paths. Consequently, we can state that the concept of providing learning objectives based on dedicated topic units is reasonable and accepted by the learners. For the mainframes2018 course, an objective was provided for learners who wanted to have a look at the course to find out whether the course is interesting to them or not. This group is of considerable size with the objective being chosen the second-most. The learning objective feature can therefore help to identify a variety of intentions.

In the related work presented before, learners had tended to change their objective during the course. With the objectives integrated into the platform, the objective can now be explicitly set and adjusted as needed. However, changes between objectives rarely happened in the examined courses. In the s4h12 course, only 2.7% of the learners changed their objective during the course while in mainframes2018 even fewer learners (0.8%) adapted it. Two possible reasons may contribute to this low rate of changes. First, learners may simply not know or remember how to change the objective as this is only described upon selecting an objective, but there is no further note or hint throughout the course yet. Another reason could be that learners change their objective for the course but do not reflect the change by selecting a new objective on the platform. A general trend can be recognized for the s4h12 course concerning the type of changes. The majority of learners switches to larger, more demanding objectives rather than between topic objectives or less demanding objectives (Table 3.10). This suggests that learners get motivated to exceed their initial intention.

Table 3.10: Descriptive Statistics for the Type of Changes of the Selected Objective

	s4h12		mainframes2018	
Objective Level	Total	Quota	Total	Quota
To Higher	45	0.763	2	0.500
Equal	4	0.068	0	0.000
To Lower	10	0.169	2	0.500

Usability and Usefulness of Learning Objectives

In addition to the A/B/n test, a survey was conducted to assess the usability and usefulness of PLOs, next to the acceptance. This further elaborates the answer to research question 2.8.

Methodology According to the well-known technology acceptance model (TAM) suggested by Davis [26], two factors have a decisive influence on the acceptance of an information system: its usability and the perceived usefulness. For this reason, the user survey particularly addressed the usefulness of the objectives itself and the usability of the selection modal to understand the influence of the chosen design and the provided information. To quantitatively measure the learners' perception of different aspects, the participants mainly had to rate their agreement with given statements on the basis of a five-point Likert scale. Because this type of questions is not diagnostic, the survey was complemented by open-ended questions to gather qualitative feedback. The conducted survey targeted all participants who took part in the preceding A/B/n test and were able to choose a learning objective, i. e., students who were assigned to the groups 2, 3, or 4. There was a total of 163 complete survey submissions from learners. The detailed figures can be found in Appendix C.3.

Analysis and Discussion The majority of the learners (77.30%, Q1) who participated in the survey also selected an objective on the platform. Most of these learners were interested in trying out the new feature (57.14%) and in the experience of choosing and following a learning objective for the course (53.97%, Q3). While the available number of learning objectives was sufficient for the majority (55.21%), some learners prefered to have even more objectives available (20.86%, Q2).

In general, the results regarding the learning objective selection, specifically the selection modal, show that the usability is perceived well (Q4). Particularly, 59.51% of the participants liked the presentation with a modal (Q4.7) and no usability issues were reported. Further, 54.61% of the learners considered the selection of an objective as useful (Q4.4) with half of them stating that it helped them to achieve their personal goals (50.92%, Q4.6). However, compared to the usability, there is a stronger variation between the rating of the participants. The usefulness was considered slightly worse, which could be related to the yet limited use of learning objectives throughout the platform. In this experiment, the selection of a learning objective supported in determining the learning path but did not relate to other activities, such as the evaluation of the learning outcome, so far. With regard to the provided information for each objective, the users agreed or even strongly agreed that it is useful (66.87%, Q5.1) and sufficient (61.35%, Q5.2) to decide on their objective for the course. While 38.27% of the learners found the time effort information helpful but did not explicitly use it for their decision, 33.33% of the participants utilized it as a decision criterion (Q6). This confirms the relevance of the provided information to allow learners to adequately choose the best-suited learning path based on personal (time) constraints.

Also, several motivational effects and an influence on the students' learning process are recognized. From the learners' perspective, the objective selection helps to become clear about the primary interest for the course (50.92%, Q5.4) and to focus on it appropriately (53.38%, Q5.5). Moreover, the learning objective motivated 49.08% of the learners to commit to the course (Q5.6) and improved their learning effectiveness (50.30%, Q5.9). An influence of the motivation to complete the learning material stands out as well since the majority of learners answered that they at least completed the material included in the objective (51.53%, Q5.7) or the objective motivated them to complete even more material than they initially intended to complete (42.95%, Q5.8).

Besides, the survey confirmed the results of the experiment with regard to the limited number of changes between objectives. Only 7.98% of the participants answered that they changed their objective during the course (Q9). Although changes have happened rarely in total, it emphasizes one of the major advantages of integrating learning objectives in the platform: it can be adapted if needed. After selecting an objective, the respective learning content is highlighted as described. Most learners who selected an objective did adhere to the course structure (26.88%), and 19.38% of the participants were focusing on the highlighted content only. Additionally, the learners were motivated by the objective to work on additional content (25.62%) and thus did not only adhere to the items being part of the objective (Q7). In total, the highlighting of the learning resources was helpful for 65.03% of the participants (Q8.1). With regard to the usability of the highlighting, the users confirmed that it is clearly distinguishable which learning items belong to the objective as well as which sections contain respective learning content.

To sum up this part of the survey, the participants were asked to rate the learning objectives concept with stars ranging from 1, being the worst, to 5, being the best. A proportion of 77.17% gave 4 or even 5 stars resulting in a mean of 4.08 with a standard deviation of 1.01 (Q11). Moreover, 64.42% would like to have learning objectives available in other courses as well (Q10). It can be concluded that the survey yielded positive results regarding the objective selection and the concept in general. The motivational component is most beneficial and could be related to improved learning outcomes, which can be examined in future studies.

3.6.5 Conclusion

This study examined how personalized learning objectives can be integrated into a MOOC platform (research question 2.7) with the aim of explicitly providing different learning paths and engaging learners in two SRL strategies: goal setting and strategic planning. For this purpose, adequate tools for facilitating these activities have been developed for the HPI MOOC Platform. First, the concept of learning objectives in MOOCs was defined considering educational and platform-related limitations. Building on that, a new service was designed and implemented allowing to flexibly create objectives for multiple use cases. These objectives can particularly cover different topic units of a course, and thus individual learning needs can be addressed better. The learners are provided with an interface for selecting learning objectives and subsequently supported by guiding the learning process with respect to the selected objective. Consequently, a learner can now follow individual learning paths while receiving guidance on the attainment of personal goals at the same time.

To provide a conclusive answer on how the new feature is perceived by the learners (research question 2.8), the concept was examined with a mixed-method approach. First, an A/B/n test in two courses analyzed the learners' acceptance. The results showed that the majority selects learning objectives in the platform and further confirm the learners' varying needs including the demand for acquisition of specific knowledge. With regard to research question 2.7, nudging learners with an objective modal while offering multiple possibilities to set an objective was identified as the best-suited approach to engage learners. Additionally, a survey was conducted to receive feedback on the perceived usefulness and usability. It revealed that the tools are well-perceived by the learners. Therefore, the goal of integrating PLOs to support students in their learning was accomplished and thus also pave the way for future research. Based on the ascertained acceptance of PLOs, the next study examines how learners perform in terms of achieving the defined objectives. Also, self-evaluation is implemented into the concept of PLOs.

3.7 Personalized Learning Objectives in Practice

In this section, we discuss a follow-up study of personalized learning objectives in MOOCs to encourage self-regulated learning, support the varying motivations and intentions of students with different social and cultural backgrounds from all over the world, and break up the one-size-fits-all approach of weekly-structured courses and the certification-based definition of success. Based on the previously well-perceived acceptance and usefulness of the concept of personal learning objectives, this study examines which learners select an objective regarding their socio-demographic and geographical background (research question 2.9), compares the general course satisfaction of students with and without a selected learning objective (research question 2.10), and investigates how successfully learners achieve their selected objectives (research question 2.11). For this purpose, a mix-method approach is chosen in which platform data and a survey are evaluated. Thereby, the existing concept, which already supports goal setting and strategic planning, is extended by the possibility of self-evaluation regarding the student's learning objective through the integration of the learner dashboard.

3.7.1 Expansion of the Concept

Building on the results of the preceding studies, the existing toolkit to support personalized learning objectives is polished and extended for the current experiment. Three features for learners are technically implemented and thus directly integrated into the HPI MOOC Platform: the learning objective selection, guidance, and evaluation.

Objective Selection

Selecting an objective is always optional to not restrict the open nature of learning in MOOCs and upset users who want to stick to the traditional learning path they are used to. Therefore, personal learning objectives were introduced as an extension and optional feature. Based on the previously examined UI alternatives in Subsection 3.6.3, an objective selection modal is shown when the user accesses the learning content for the first time. This modal can be dismissed without selecting an objective and it will never appear automatically again to avoid learner frustration. If the user later decides to select an objective the modal can be opened again from an infobox displayed at the top of each learning item page or from the progress page. Here, the currently selected objective is displayed and can be changed at any time. The infobox disappears if an objective was selected or if it was dismissed by the user. Previous evaluations showed that it is reasonable to explicitly encourage learners to select an objective instead of relying on its discovery by learners. Additionally, offering objective selection at different places was recommended.

The selection modal was chosen as an interface to avoid a complete context switch from the learning process for the user and to reuse it at different pages. The modal provides information on how the learning process is affected when an objective was selected. Every objective provides a short title and expandable details by clicking on it to make an informed decision. Next to a more comprehensive description of the objective also the estimated time effort accumulated for each learning item type and the whole objective is displayed. This also enables learners to compare the different objectives better. Currently, only one objective can be selected per course. After selecting an objective, a short confirmation is shown with additional explanations on how the user is guided through the learning items of the chosen objective.

Objective Guidance

Personal learning objectives provide the opportunity to open up the traditional course structure in MOOCs. To quickly identify the content which is part of a selected objective, blue triangles were implemented as highlights in the course navigation, as shown in Subsection 3.6.3. Also, a tooltip is provided when hovering the navigation items. This enables learners to see where to start with their objective and which content they should focus on. Since the learning items of an objective can be placed in multiple weeks of a course, the left-hand section navigation is decorated with the same blue triangles. As shown, the regular course structure is still maintained and the implemented approach of guidance does not restrict the user from accessing the other content. This also enables learners to do more than initially intended and exceed their original objective.

Objective Evaluation

Next to selecting and working on a learning objective, a user needs to constantly evaluate the progress and achievement. Therefore, we adapt the platform's progress overview on the learner dashboard (Figure 3.19) based on the work in Section 3.3 and Section 3.4. At the top of the progress overview, the overall course progress summary is complemented by the overall learning objective progress summary as it becomes the main focus of learning. The objectives for the completion and the exploration of a course do not differ from the usual course progression. Hence, in such cases only the objective details are added to raise awareness for the selected personal objective. For the objectives targeting the completion of a specific topic not all course content is relevant, and therefore two more indicators are shown: the achieved points and the percentage of items visited being part of the objective. This progress is depicted using circular progress bars and thus visually distinguished from the more detailed section progresses below. If the material

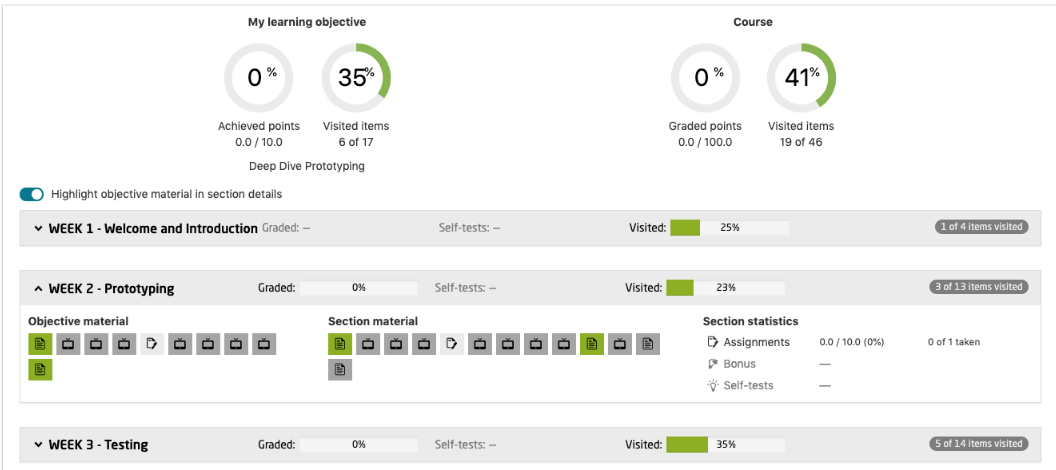


Figure 3.19: The Learning Objective and Course Progress Overview on the Learner Dashboard

for a selected objective is finalized and the student can still achieve it, this is stated in the presented summary aiming to motivate the learner to continue with the course and actively work towards the achievement of the objective.

Each section progress can be expanded and the objective material can be highlighted with a checkbox above. By activating this checkbox, the learning items which are part of the selected objective are grouped separately next to the regular material. This allows learners to follow their objective but also to discover the other course material. Each learning item is visualized as a rectangle and colored if visited or partially filled if not all points were achieved. This visualization was introduced to distinguish between different states of visited and completed content. All in all, this overview supports to raise awareness and allow learners to self-evaluate their progress. Below that part, the currently selected objective is displayed and can be changed.

3.7.2 Evaluation

The following subsection presents the evaluation of the three remaining research questions regarding personalized learning objectives. To investigate them, it is divided into an analysis of the learners' socio-demographic and geographical background, a post-course survey about the learners' satisfaction, and an analysis of the achievement rates of learning objectives.

Sample Courses

We examined two courses of openHPI. The first course was "Introduction to Successful Remote Teamwork" (international-teams2019[18]) and presented benefits and risks of driving a virtual team culture and how guided remote work leads to success, as well as intercultural competences as a key factor of interaction and communication. It was held in English and was running from October 2, 2019, until October 30, 2019, starting with 2,327 enrollments. The course was structured into four weeks and graded with a final exam (55% of all points) and one team peer assessment (45% of all points). An RoA was gained by about 20% of all 'shows at middle' by earning more than 40% of all graded points. A CoP was achieved by about 34% of all 'shows at middle' by completing at least 50% of the course material. All numbers are listed in Table 3.11 as well. The teaching team defined the following objectives:

1. *Passive Participation with Certificate*. This objective comprised all course material including the final exam to gain a Record of Achievement if the learner got most of the exam right.
2. *Active Participation with Certificate*. This objective comprised all course material including the final exam to gain a Record of Achievement if the learner got most of the exam right. Additionally, the learner was encouraged to contribute to discussions or at least follow them.
3. *Deep Dive Virtual Teamwork with Certificate*. This objective comprised all course material including the participation in discussions, the final exam, and the team peer assessment to gain a Record of Achievement.
4. *Peek In and Explore*. This objective left it up to the learner to study the material and to receive a Confirmation of Participation.

[18] https://open.hpi.de/courses/international-teams2019/

Table 3.11: Enrollments and Certificates in Sample Courses with Personalized Learning Objectives

Course	Enrollments At Start	Enrollments At Middle	Enrollments At End	Shows At Middle	Records of Achievement	Confirmations of Participation
prototype2019	3,029	3,356	3,533	1,568 (46.72%)	250 (15.94%)	626 (39.92%)
international-teams2019	2,327	2,778	2,991	1,074 (35.91%)	212 (19.74%)	370 (34.45%)

The second course on the topic of "Human-Centered Design: Building and Testing Prototypes" (prototype2019[19]) covered different task-based approaches to turn an idea into a simple prototype, set up a testing scenario, and collect feedback. The course was held in English and was running from August 28, 2019, until October 10, 2019, starting with 3,029 enrollments. It was structured into four weeks and graded with three exercises (40% of all points) and one peer assessment (60% of all points). An RoA was gained by about 16% of all 'shows at middle' by earning more than 50% of all graded points. A CoP was achieved by about 40% of all 'shows at middle' by completing at least 50% of the course material. More detailed numbers can be seen in Table 3.11. The learning objectives in this course built on user behavior observations from previous Design Thinking MOOCs: the research team observed participants who only explored partial modules of the course or browsed course contents for educational material to download. The teaching team offered the following six learning objectives:

1. *Complete Course Experience*. This objective comprised all course material including the graded exercises and the peer assessment to gain a Record of Achievement.
2. *Explore*. This objective comprised all introductory material about design thinking, prototyping, and testing. Following the objective was sufficient to gain a Confirmation of Participation.
3. *Deep Dive Prototyping*. This objective focused only on content about prototyping.
4. *Deep Dive Testing*. This objective focused only on content about testing.
5. *Material Collector*. This objective highlighted the material items for users who were mainly interested in collecting resources and templates.
6. *Inspirational Trip*. Learners who did not know whether or not the course is interesting to them could choose this objective to take a look at the course.

Methodology

To investigate research question 2.9 and 2.11, the platform's data of the two sample courses were analyzed after the courses had ended and the final results and certificates were released. Therefore, course reports were used to export information from the platform about each enrollment's sociodemographic profile data and metrics about the learning behavior and course completion. The reports were enriched with data about the selected objectives and their achievement. All exported data were pseudonymized and analyzed with external tools afterward. The enrollments are filtered by 'shows at middle'. Users who have never showed up for the course or never had the chance to

[19] `https://open.hpi.de/courses/prototype2019/`

achieve a graded certificate are excluded from the study. Gaining a certificate is the traditional way to measure course success and also the most demanding learning objective in both courses. Therefore, only users who have had a realistic chance to do so are of interest for the study and form the total population in this scope. Since the users could choose an objective voluntarily, this was not a controlled experiment with a control group but an authentic and real-world learning experience. Thus, we compare users with a selected learning objective with the total course population.

To gather information on the students' satisfaction and address research question 2.10, a post-course survey was conducted in week four of the prototype2019 course, in which students could participate voluntarily. Unfortunately, the survey could not be conducted in the international-teams2019 course as well. There was a total of 279 complete submissions regarding the assessed questions, 163 from students without a selected learning objective and 116 from students with a selected learning objective. The first four questions could be answered with the use of a Likert scale including answer options from 'not satisfied at all' (1) to 'absolutely satisfied' (10):

1. Please rate this MOOC by indicating how satisfied you are with the overall course.
2. How satisfied are you with the quality of the content presented in the course?
3. How satisfied are you with the length of the course?
4. How satisfied are you with the openHPI learning platform?

After that there was a single-choice question with the answer options 'no' (1) and 'yes' (2):

5. Were your personal learning expectations met?

Based on the numerical value of each answer option, we compare both user groups for statistically significant differences utilizing the nonparametric Mann-Whitney U test for two independent samples. We also assess the practical relevance of the descriptive statistics, based on the authors' long-term experience with the operation of several MOOC platforms and courses.

Socio-Demographic and Geographical Background

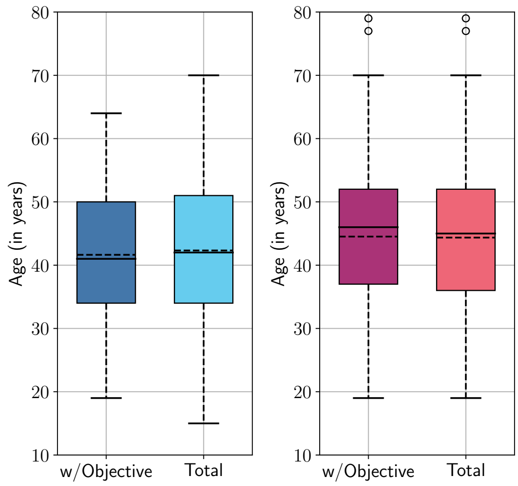
The students' socio-demographic data is based on their platform profile. They can provide this information voluntarily. Therefore, this data is not available for each student and missing entries are excluded. For the prototype2019 course about 33% have provided this data and for the international-teams2019 course about 44%.

Table 3.12 displays the gender distribution in both courses. It can be seen that roughly one quarter in both populations is female and the rest is male. This also applies for the cohorts of users who selected an objective. Again, we cannot identify a practically significant difference in this characteristic.

Table 3.12: Gender Distribution of Learners

	With Objective		Total	
Course	Female	Male	Female	Male
prototype2019	27%	73%	29%	71%
international-teams2019	26%	74%	25%	75%

In Figure 3.20 the age distribution is displayed for both course populations and the students who selected a learning objective. It can be seen that there is no practically significant difference of the learners' age. The mean age of users in the prototype2019 course is 42.32 with a median of 42 years and the mean age of users with an objective in this course is 41.64 with a median of 41 years. The mean age of users in the international-teams2019 course is 44.36 with a median of 45 years and the mean age of users with an objective in this course is 44.52 with a median of 45 years. Figure 3.21 shows that more than 80% of users attended university and the majority of them gained a Master's degree. The results highlight that also the educational background of users does not differ noteworthy between the two courses and users with an objective.



(a) prototype2019 (b) international-teams2019

Figure 3.20: Age Distribution of Learners

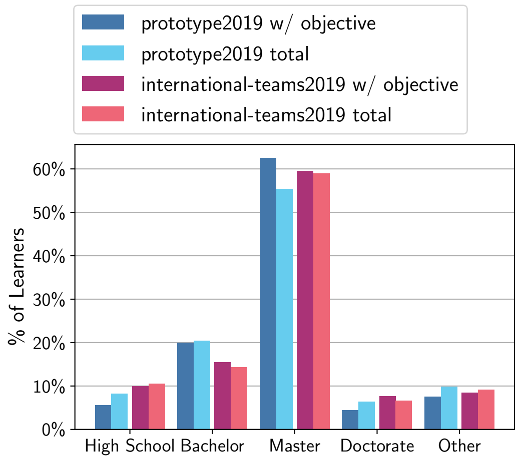


Figure 3.21: Highest Degree of Learners

As the last socio-demographic characteristic, we examine the career status (Figure 3.22). Again, there are very similar results with no practically significant differences. Around 80% of the learners were professionals and the remaining were students, researchers, teachers, and others. We also examine the geographical background of each user (Figure 3.23). Therefore, the learner's IP address is mapped to a location for each action and the country with the highest frequency is picked. With this automated process, the information is available for all learners and no self-reported data is needed. It has to be noted that this variable reflects the country where most of the learning activities took place but not the nationality of a user (although there is probably a strong correlation). Even though both courses were offered in English, most users accessed them from Germany (up to 70%). This is mainly because the platform originates from Germany and is best known there. Users were further distributed among other German-speaking countries, such as Switzerland and Austria, followed by more populous nations such as the USA, India, and other countries. However, again, the distributions between courses and users with an objective are very similar and there are no practically significant differences.

In summary, the examined learners are on average well-educated, working men in their mid-40s which is very typical for MOOCs and diversity is still a large issue [46, 67]. Concerning research question 2.9, we did not identify any practically significant differences between students with selected learning objectives and the total course population regarding their age, gender, degree, career status, or geographical location. Of course, this statement is limited to our case study and a larger sample from other contexts is required to assess its external validity.

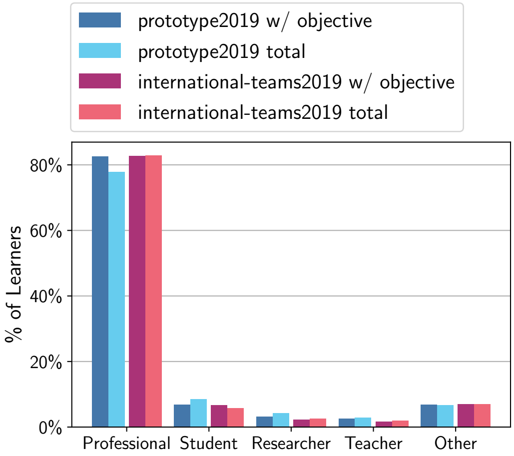


Figure 3.22: Career Status of Learners

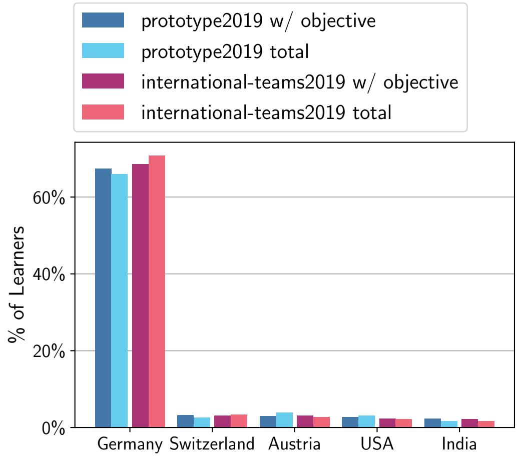


Figure 3.23: Geographical Location of Learners

Course Satisfaction

Before we discuss the results of the statistical analysis of the course satisfaction, we first review the learning success of the survey respondents. Many users drop out of a MOOC during its runtime for a variety of reasons. Therefore, when surveys are conducted at the end of a course, it is usually only the most engaged and successful learners who have made it to this point who participate. Therefore, we want to assess the extent to which a survivorship bias may influences the results of the survey.

Table 3.13: Achieved Certificates (prototype2019)

Cohort	Count	Records of Achievement	Confirmations of Participation
Shows at Middle	1568	250 (15.94%)	626 (39.92%)
Survey Respondents	279	214 (76.70%)	279 (100.00%)

Table 3.13 shows the number of certificates achieved by the 'shows at middle' of the entire prototype2019 course, i. e., the users who had a realistic chance of receiving a graded certificate, and the survey respondents. It can be seen that the survey respondents have gained a much higher number of Records of Achievement (76.70%) than the 'shows at middle' (15.94%). In addition, all survey respondents achieved a Confirmation of Participation. Only 39.91% of the 'shows at middle' achieved this. Table 3.14 displays the selected learning objectives and achievement rates of the 'shows at middle' and survey respondents. In both cohorts the clearly most frequently selected learning objective is the 'Complete Course Experience' which includes the completion

of a Record of Achievement (71.88% and 83.63%). Also in these subsets, the survey respondents reached the learning objective far more often (83.51%) than the 'shows at middle' (28.57%). The second most frequently selected learning objective is 'Explore' which includes the achievement of a Confirmation of Participation (14.73% and 11.21%). All survey respondents completed this objective, but only 46.97% of the 'shows at middle'. The other topic-based learning objectives were selected far less frequently, especially by the cohort of survey respondents. We discuss more details of the achievement rates of learning objectives in the next part of this evaluation.

Table 3.14: Selected and Achieved Learning Objectives (prototype2019)

		Shows at Middle			Survey Respondents		
Objective	Type	N	Quota	Achieved	N	Quota	Achieved
Complete Course Experience	RoA	322	71.88%	28.57%	97	83.62%	83.51%
Explore	CoP	66	14.73%	46.97%	13	11.21%	100.00%
Deep Dive Prototyping	Topic	31	6.92%	19.35%	3	2.59%	100.00%
Deep Dive Testing	Topic	4	0.89%	0.00%	0	0.00%	-
Material Collector	Topic	3	0.67%	0.00%	0	0.00%	-
Inspirational Trip	Topic	22	4.91%	18.18%	3	2.59%	66.67%

The examination of the results in Table 3.13 and Table 3.14 has shown that the survey respondents have clearly better learning outcomes than the course population, measured by the 'shows at middle', both in terms of the traditional completion with a certificate and in achieving personal learning objectives. It therefore has to be assumed that the answers of the survey respondents are subject to a survivorship bias and that this shifts the overall course satisfaction into a more positive direction.

Table 3.15: Descriptive and Inferential Statistics for Survey Respondents With and Without a Selected Learning Objective (prototype2019)

	With Objective			Without Objective			Mann-Whitney U	
Question	N	Mean	Std.Dev.	N	Mean	Std.Dev.	U	p-value
1	116	8.422	1.610	163	8.441	1.461	9342.5	0.862
2	116	8.500	1.568	163	8.404	1.573	9115.0	0.597
3	116	8.172	2.022	163	8.404	1.780	8887.0	0.378
4	116	8.474	1.917	163	8.582	1.756	9242.0	0.739
5	116	1.939	0.239	163	1.914	0.281	9212.5	0.426

The results of the questions examined in the survey are shown in Table 3.15. The respondents are divided into users with ($N = 116$) and without ($N = 163$) a selected learning objective. It can be seen that the overall course satisfaction (Question 1) is almost the same for both cohorts (8.422 and 8.441). The satisfaction with the quality of the course content (Question 2) was slightly better perceived by users with a selected learning objective (8.5 and 8.404). This trend may be due to the fact that relevant content is better highlighted for users with learning objectives. Interestingly, users without learning objectives are a little more satisfied with the length of the course (8.172 and 8.404) which was asked in Question 3. We assume that users who choose a learning objective are more likely to have a specific focus in the course and are not interested in the whole content, which is supported by the concept of personalized learning objectives. In Question 4 users were asked for their overall satisfaction with the learning platform and both cohorts show very similar positive results (8.474 and 8.582). The same applies to Question 5, in which users were asked whether their personal learning expectations were met (1.939 and 1.914).

Overall, the generally very positive results show no statistically significant differences between the two cohorts based on the calculated p-values. Furthermore, we cannot derive any practical relevance from the very small differences. It also has to be noted that the very positive results are distorted by a proven survivorship bias. Research question 2.10 can therefore be answered to the effect that in the Design Thinking MOOC studied, students with a selected learning objective are no more, but also no less satisfied with the course than students without a selected learning objective.

Achievement of Learning Objectives

The total objective selection rates for both sample courses are presented in Table 3.16. The big difference between 28.57% and 63.87% of users who selected a learning objective is probably due to the fact that objectives are defined individually for each course by different teaching teams. Therefore, they differ in their complexity and in the way they are formulated. Furthermore, the diverse course contents may attract users from different industries and backgrounds. Unfortunately, this information is not included in the socio-demographic profile data. Compared with the objective selection rates from the two courses of the previous study (49.4% and 58.7%) in Subsection 3.6.4, it can be stated that a notable portion of learners sets a personal learning objective and it further confirms the acceptance of the concept. However, the objective selection rates can differ largely between courses. The number of changes of learning objectives by users are negligibly small, which was also shown in Subsection 3.6.4.

Table 3.16: Total Selected Learning Objectives

Course	Shows At Middle	With Objective	Quota
prototype2019	1,568	448	28.57%
international-teams2019	1,074	686	63.87%

Table 3.17 and Table 3.18 display in detail the selected objectives of each course, as well as their achievement and exceeding rates. We also defined a criterion for each objective by which it could be measured whether a goal was exceeded. Objectives that already included the Record of Achievement could not be exceeded. Objectives that included the Confirmation of Participation (some topic objectives included that in the textual description) could be exceeded by gaining an RoA. All other topic objectives, which only included a minor subset of learning items, could be exceeded by achieving a CoP. It can be seen that in both courses the most frequently selected learning objectives include an RoA, which indicates that a large amount of users is still interested in completing the course with a graded certificate. This is reasonable considering that the courses were also primarily designed with this intention. However, this is not the case for a notable amount

of learners who were not interested in gaining an RoA (28.12% and 16.33%). This is also supported by the numbers from the previous study (34.3% and 44.1%) confirming our assumption that learners enroll for courses with varying outcome intentions. For the prototype2019 course, the two objectives 'Deep Dive Testing' and 'Material Collector' have almost never been selected. This was unexpected since at least the other deep dive objective was chosen more frequently. The reasoning for these outliers needs to be further investigated with qualitative user feedback.

Table 3.17: Selected Objectives with Achievement and Exceeding Rates (prototype2019)

Objective	Type	Selected	Quota	Achieved	Exceeded	Criterion
Complete Course Experience	RoA	322	71.88%	28.57%	-	-
Explore	CoP	66	14.73%	46.97%	6.06%	RoA
Deep Dive Prototyping	Topic	31	6.92%	19.35%	9.68%	RoA
Deep Dive Testing	Topic	4	0.89%	0.00%	0.00%	RoA
Material Collector	Topic	3	0.67%	0.00%	0.00%	CoP
Inspirational Trip	Topic	22	4.91%	18.18%	18.18%	CoP

Table 3.18: Selected Objectives with Achievement and Exceeding Rates (international-teams2019)

Objective	Type	Selected	Quota	Achieved	Exceeded	Criterion
Passive Participation	RoA	249	36.30%	23.29%	-	-
Active Participation	RoA	174	25.36%	33.91%	-	-
Deep Dive Virtual Teamwork	RoA	151	22.01%	41.72%	-	-
Peek In and Explore	CoP	112	16.33%	23.21%	8.93%	RoA

In this study, learners had the opportunity to evaluate the achievement of their learning objective on the new progress overview on the learner dashboard for the first time (Figure 3.19). This is made possible by the automatic calculation of the objective progress within the platform also allowing us to evaluate the achievement rates. From our perspective, with the experience of many years of operating multiple MOOC platforms, these figures ranging from 18.18% up to 46.97% are considered a success. Also, between 6.06% and 18.18% of the learners even exceeded their objectives, which indicates an increase in motivation during the course. However, to the best of our knowledge, there are currently no comparable figures as this approach of success in MOOCs is new. Only the certification rates for objectives that included a Record of Achievement can be compared with the traditional approach (Table 3.19) since gaining a graded certificate is the commonly assumed course outcome to date. For the prototype2019 course, we see an increased certification

Table 3.19: Traditional vs. Selected Objective Certification Rates

Course	Shows	Traditional RoAs		Selected	Objective RoAs	
	At Middle	*N*	Quota	RoA Objective	*N*	Quota
prototype2019	1,568	250	15.94%	322	92	28.57%
international-teams2019	1,074	212	19.74%	574	180	31.36%

rate of 12.63%, from 15.94% of gained Records of Achievement for the total course population to 28.57% of users who selected objectives including this certificate. For the international-teams2019 course, we see an increased certification rate of 11.62%, from 19.74% of gained Records of Achievement for the total course population to 31.36% of users who selected objectives including this certificate. We consider both rates as a practically significant improvement. However, these rates relate to different total quantities and therefore do not reflect an absolute increase in the number of gained certificates. Nevertheless, they demonstrate that this objective achievement-based method is more reasonable for calculating completion rates in MOOCs than the traditional approach.

All in all, we have observed varying objective selection rates, probably due to different formulations by different teaching teams and target groups. Most users tend to select objectives that include a graded certificate, but also a considerable number of learners select objectives with less effort, covering only parts of a course. Regarding research question 2.11, about one-fifth to half of the learners achieve their learning objectives and a notable amount of them even exceeds them. We were also able to compare the certification rates of the total course population with the users who selected an objective which leads to a graded certificate. We observed a practically significant improvement.

3.7.3 Conclusion

In this section, we presented a continuative study of personalized learning objectives in MOOCs to better support self-regulated learning and incorporated the three strategies goal setting, strategic planning, and self-evaluation with technical means. These are crucial skills for students' achievement and success in online learning environments with little support and guidance like MOOCs. Having explored the students' acceptance and usefulness of the concept in the last section, the overall positive results prompted us to further investigate which learners select an objective, and how successful they complete objectives. Regarding the learners' socio-demographic and geographical background, i. e., their age, gender, degree, career status, or geographical location, we did not identify any practically significant difference between students with selected learning objectives and the total course population (research question 2.9).

Furthermore, we compared the course satisfaction of students with and without a selected learning objective with self-reported data from a post-course survey. We found no statistically significant differences and no practical relevance. Therefore, students with a selected learning objective are no more, but also no less satisfied with the course than students without a selected learning objective (research question 2.10). It is interesting to note that the provided tools do not seem to have any impact on the general course satisfaction, but usefulness and achievement rates are perceived and influenced positively. To investigate causality, further qualitative studies are necessary.

However, we identified promising objective achievement rates. Additionally, we observed a practically significant improvement of the certification rates comparing the total course population and students who selected an objective that included a graded certificate (research question 2.11), which provides further evidence that our concept of personalized learning objectives in combination with a learning analytics dashboard fulfills its intended purpose. The selection and achievement of learning objectives enable users to pursue their individual intention and motivation to enroll in an open online course, whereby they can evaluate their learning progression by means of data-driven insights. Moreover, this concept also contributes to further detach the definition of learning success in this open format from the exclusive achievement of certificates.

3.8 Summary

This chapter addressed the question of how data-driven insights can be used to support learning in MOOCs. Due to the massive number of students in such online courses, individual guidance often cannot be provided. Therefore, it is crucial for the learner's success to set goals autonomously, plan them strategically, evaluate their progression, and adjust the own learning behavior if necessary. However, many students do not self-regulate their learning because these skills need to be learned and practised as well. Here, technical means can provide support. For this purpose, we developed a learning analytics dashboard for learners, and also elaborated and integrated the concept of personalized learning objectives into the HPI MOOC Platform. In several iterations, we investigated various aspects and research questions. Thereby, a wide range of the platform's learning analytics capabilities was used.

First, we expanded the existing progress page of the platform into a dashboard. For this purpose, requirements were derived from literature on the one hand and from an ideation session with experts on the other hand. In the latter, six categories emerged to be considered in the development of visualizations and widgets: an improved progress overview, the invested learning time vs. the outcome, the time needed for the remaining course material, a more fine-grained performance evaluation, opportunities for social comparison, and call-to-actions. These, along with the requirements and limitations of technical feasibility, were incorporated into the implementation process of the dashboard. The following components were developed: a revised progress overview, the estimated time for course material, next course dates, the students' forum activity, information on the self-test and assignment performance, repetition suggestions, the achieved points over time, the time spent on quizzes, the timeliness of submissions, the visited items over time, the learning activity per hour per weekday, and session statistics. Also, empty states for the widgets were implemented to encourage the learner to reflect and adjust their learning behavior if necessary. Furthermore, the entire technical foundation of the page was replaced in order to load and render the various components independently of each other. In a first study, a survey was conducted to investigate whether the dashboard is perceived as useful and accepted, as well as which components are most valued by the learners. The majority of participants agreed or strongly agreed that it is easy to use and extremely useful. The learners' acceptance is therefore clearly present. The most valued widgets on the dashboard were the progress overview, the course dates, the achieved points over time, the quiz performance, and repetition suggestions. Other ideas for improvement were also stated.

In the next iteration, we first incorporated the learners' feedback. Therefore, some dashboard components were rearranged, more explanations were added, and the session statistics were removed. We then conducted an A/B/n test to determine whether the completion rates were statistically significantly different when learners used the old progress page, the new progress overview, or the complete revised dashboard. However, the results clearly showed that there are no significant differences in the completion rates of learners concerning the use of the three dashboard variants and thus all are sufficient to achieve a graded certificate. Nevertheless, one of our main aims is to support self-regulated learning, which we explored further with a survey. The responses revealed that in particular the strategy of self-evaluation is encouraged by the stimulation of awareness and reflection of one's own learning situation and behavior. Besides, strategic planning is also partly supported by stimulating the adaptation of the learning behavior when required.

The next step was the incorporation of the strategy of goal setting. Therefore, we first examined the status quo of how learners in MOOCs set and achieve learning objectives. For this purpose, self-reported pre-course surveys of the HPI MOOC Platform were analyzed and the results were compared with similar studies. More than two-thirds of users selected a learning objective and of those, only about one quarter were interested in a graded certificate. The majority was therefore mainly interested in learning the course content, regardless of a formal record. Based on the analysis of platform data, we found that approximately half of the learners achieved or exceeded their goal, but the precise ratios are very course-specific. The presented related studies also showed very diverse objective achievement rates. On the one hand, this is probably due to the different course characteristics in terms of instruction design, examination modalities, difficulty, and required effort. On the other hand, it is also caused by the imprecise and inadequately comparable analysis methods utilizing pre-course and post-course surveys. Therefore, we elaborated a concept for personalized learning objectives, which incorporates goal setting, strategic planning, progress evaluation, and goal achievement as technical platform features.

Thus, we developed a learning objectives model that supports course-level objectives including certificates, thematic units, and course exploration. This model was implemented within a new service for the service-oriented architecture of the HPI MOOC Platform. To set an objective, different UI alternatives were developed that can be used at different places and times in a course, of which the most accepted were examined later. The selection of an objective is accomplished with a modal that lists all objectives with descriptions and the required time effort. It also outlines how the objective impacts the learning process and afterward a confirmation explains the objective guidance, for which course sections and items are marked that are part of the selected objective. Also, the objective can be changed at any time on the progress page. We conducted an A/B/n test to assess the acceptance of the implemented prototype. Of all learners, about one-third to one-half selected a learning objective. The UI pattern that performed best was to automatically open the modal once the first time a learning item was visited, and also to display an infobox on each item page that can be used to manually open the modal. Over one-third of learners selected a goal that did not include a certificate. Changes to the objective after initial selection were rarely made. In addition, another survey was conducted to determine the usefulness and usability of the prototype. Positive results were measured here, especially with regard to motivating aspects of the concept.

In the last step, the concept was refined based on the results of the previous test. The new progress overview of the learner dashboard was incorporated to be able to track the progress of the selected objective. Thus, the three self-regulated learning strategies goal setting, strategic planning, and self-evaluation were combined and technically supported within the platform, whereby the prototypes realized one practical approach out of many. For the last evaluation, we explored three research questions. First, we examined the sociodemographic and geographic background of learners with selected learning objectives. The profile data evaluated for this purpose did not reveal any practically significant differences between learners with selected objectives and the total course population regarding their age, gender, degree, career status, or geographical location. Second, we used a post-course survey to examine whether the general course satisfaction differs between learners with and without a selected objective. The results were very positive for both cohorts but did not show any statistically significant differences. Last, we examined the achievement of learning objectives. The varying selection rates and types of selected learning objectives support the conclusions of the last study. Depending on the objective and course, about one-fifth to one-half of the learning objectives were achieved. Compared to conventional completion rates in MOOCs, these figures are a noticeable success. In addition, even approximately six to eighteen percent of the learning objectives were exceeded. This had not been measurable before and the learning success could not be determined in such detail. However, this also makes it difficult to compare the rates with existing figures. We were therefore only able to make a direct comparison of learners who selected an objective that included a graded certificate with traditional certification rates of the total course population. Here, we found increased certification rates of about 11.6% to 12.6% when a learning objective was selected. We consider this a practically significant difference which emphasizes that the concept is fulfilling its purpose of supporting self-regulated learning and is having a positive impact on students' achievement.

All in all, we demonstrated that technology incorporating data-driven insights can support learning and self-regulation in MOOCs. A LAD for learners served as a central component, allowing them to additionally select and pursue a learning objective while evaluating their progression. Thereby, we also showed the advantages and limitations of the technical integration in an authentic MOOC environment.

4 Supporting Teaching with Data-Driven Insights

This chapter shows how a learning analytics dashboard is utilized to support teaching in MOOCs. The context-specific origin of a first prototype is explained, on the basis of which requirements are defined through teacher interviews. Technical and functional improvements are implemented and the usefulness of the integration of web analytics is investigated. Afterward, the usage and perception of the dashboard is examined.

4.1 Introduction

Already in traditional classrooms, monitoring their students' learning is a vital method for teachers to deliver quality education effectively [20, 86]. With the advent of MOOCs, classrooms became purely digital, filled with thousands to hundreds of thousands of students, most of whom communicate and interact asynchronously with each other and teaching staff through video lectures, forum discussions, and feedback surveys. The challenge of monitoring this tremendous amount of students in MOOCs is evident and cannot be achieved with human resources alone, but is even more important due to the lack of personal guidance and feedback in online and distance learning [42]. The technology that enabled these online learning environments in the first place is therefore also in demand to support teachers with these shortcomings. Hence, this chapter examines the following research question:

Research Question 3: How can data-driven insights support teaching in MOOCs?

One of the main approaches of learning analytics is to provide visual information [125]—e. g., to reveal data-driven insights for teachers into their students' learning and to support them in decision-making. Humans are capable to process large amounts of data if they are presented visually and meaningfully [72, 134] and thus empower to make better use of human intelligence [6]. Therefore, learning analytics dashboards are used to process and illustrate learning data with various visualization techniques [92]. After providing such learning analytics capabilities for the HPI MOOC Platform and utilizing them to support learning in MOOCs, it also raises the possibility to assist teachers. A dashboard had been created for teachers whose needs first evolved from the day-to-day use of the platform. This resulted from the platform's different deployment contexts and the involved stakeholders as elaborated in Subsection 1.4.2. To formalize the dashboard's requirements we explore the following sub-question in this chapter:

Research Question 3.1: Which requirements do teaching teams have for a learning analytics dashboard?

Based on the identified shortcomings, the existing dashboard prototype is technically revised. In particular, the web analytics capabilities are utilized which were integrated into the platform as introduced in Section 2.4. To investigate the added value for teachers of this extension, another sub-question is studied:

Research Question 3.2: Can web analytics methods improve the usefulness of the dashboard for teaching teams?

After that, we address functional improvements of the dashboard. The implementation allows to understand how beneficial the LAD for teachers is during the development, facilitation, and evaluation of courses. Therefore, the following two research questions are formulated, the answers to which also provide further evidence relevant to the overall research field in the form of a real-world case study:

Research Question 3.3: How are the dashboard and statistics used by teaching teams?

Research Question 3.4: How are the dashboard and statistics perceived by teaching teams?

To investigate and answer the main research question 3, the first two derived sub-questions are examined in Section 4.2, and the latter two are addressed in Section 4.3. After that, Section 4.4 provides a summary of the chapter.

4.2 Towards a Learning Analytics Dashboard for Teachers

After the early days and first courses on the platform, it quickly became clear that teaching teams need data-driven insights into their courses, for example, to identify and respond to problems, or evaluate the courses after completion. Therefore, this section shows the transition from an initial dashboard prototype that emerged from the platform's daily operation to a scientifically accompanied approach in order to systematically review and support the process of further development. Thereby, the requirements of teaching teams for a learning analytics dashboard are elaborated (research question 3.1) and first technical improvements are examined. In particular, the use of web analytics methods is addressed (research question 3.2).

4.2.1 The Origin of the First Dashboard for Teachers

In the early and more turbulent years of the platform, many features emerged from the lessons learned in the day-to-day business, especially when the first partners deployed the platform under their brand. In order to support and evaluate courses, as well as for presentations, news, and to answer management inquiries, requests were repeatedly made to the development team to provide specific figures and data. Initially, these requests were executed manually, mostly as database queries on the platform's servers. Gradually, repetitive requests were translated into features and thus made available in self-service via the platform itself. Various reports and dashboards were thus provided, including a prototype of a course dashboard for teaching teams.

This prototype includes high-level data on the number of learners and 'no-shows', the forum activity, the number of certificates achieved after completing a course, as well as information on the geographic distribution of learners and what types of devices they use. In addition, other statistics are available, such as how often a course has been shared on social networks.

However, a systematic development process is lacking and requirements and shortcomings of the dashboard are unclear. Therefore, in the following, we study related work and elaborate issues of the prototype to revise the concept and increase its usefulness.

4.2.2 Related Work

The term 'dashboard' refers to a "visual display of the most important information needed to achieve one or more objectives; consolidated and arranged on a single screen so the information can be monitored at a glance" according to Few [39]. It is thus a reporting mechanism that aggregates and visualizes metrics and KPIs to support decision-making by uncovering and communicating contextual insights [43]. In the educational technology domain, LADs are often used to visualize learning traces for learners and teachers to increase their motivation, autonomy, effectiveness, and efficiency. For teachers, most of these dashboards are developed to raise awareness of their students' activity, to reflect on their teaching practice, to find students at risk or isolated students, and to adapt their teaching behavior [160]. In the development of LADs and other LA tools, various disciplines must be taken into account, i. e., software engineering, human-computer interaction, computer graphics, educational design, and psychology [84].

One of the major challenges for dashboards is to present data in a way that can be interpreted by teachers and lead to actionable insights [134]. Studies showed that teachers often find it easy to work with dashboards, but have difficulty turning LA-based recommendations into actions for students in need [51]. Rienties et al. [108] examined that the majority of teachers are looking for training and support in the use of LA tools. In addition, the general acceptance of technology correlates positively with the satisfaction of such training, hence extra support is necessary for individuals with a lower acceptance rate. Besides general usability, Isaias et al. [55] also propose a training process based on the cycle: awareness, training, and monitoring. To improve awareness, they suggest the presentation of successful case studies and peer guidance between teachers. The training has to teach the everyday use of dashboards and how they can be utilized to improve their teaching process. In the monitoring phase, teachers have to be provided with support during their actual use of dashboards when conducting courses.

According to a study by Stephens-Martinez et al. [150], teachers especially value dashboard visualizations about the students' performance, activity patterns—e. g., which materials students use—as well as "what students have to say"—i. e., the forum behavior and survey feedback. However, it has to be noted that only a few students in MOOCs are active in the forum and that they are therefore not representative of a course's entire learning community. Also, a lot of MOOC instructors perceive demographic and geographical insights to be very motivating [41]. For the development of LADs, a user-centered design approach is recommended, since an iterative process allows to better involve the actual target group in continuous cycles of design-implementation-evaluation [72, 164].

4.2.3 Elaboration of the Concept

In the following, we describe issues with the dashboard prototype that are identified through expert interviews. Based on the gained insights, we elaborate functional and technical requirements, of which the latter are incorporated into a revised concept at first.

Issues of the Prototype

Several issues with the existing prototype were identified by conducting semi-structured interviews with relevant experts. Six employees of openSAP were interviewed about usage scenarios of the dashboard. The participants hold different occupational roles in context of openSAP. Three of them are in charge of certain courses as members of the corresponding teaching teams while two hold the role of the platform owner. The last person has experience in both roles. Despite the small number of available interviewees, the gained insights are highly relevant. The respondents are experts in their field of duty, who work with the dashboard on a daily basis. Besides, openSAP is a professional MOOC platform with a notable amount of users. Consequently, the views and opinions of the interviewees are considerable in this context. The employees were asked for which purposes they use the dashboard in their daily work and how they accomplish these tasks. In this context, special attention was paid to identify parts of the dashboard that are essential and those that are not used at all by the individual persons. Additionally, the interviewees were asked for technical and conceptual issues as well as suggestions for improvements for the existing solutions.

The interviewees were interested in the general performance of a course which they derive from the shown KPIs. Most important in this context are the number of enrollments and amount of 'no-shows' as well as the number of learners that received an RoA at the end of the course. Also, they are responsible for identifying problems of learners and support them in finding a solution. This includes answering learners' questions in the forum and examining statistics about quizzes to detect questions that could be erroneous or too difficult. For this purpose, the dashboard is of little avail. It is used only to receive an overview of how many new posts there are on the forum.

One problem, which was mentioned by all interviewees, is the pages' performance. Especially when loading the dashboard, it takes a lot of time until the page is eventually shown in the browser. In addition, the dashboard is usually visited frequently, which reinforces the issue. The reason for these long loading times is that the page is not rendered until all required LA data is loaded and visualized metrics and statistics are calculated.

Three of the interviewees criticized the dashboard to be cluttered. The page contains many different visualizations and the majority of them is not relevant for all stakeholders. As a consequence, users may scroll over a number of components until they reach the visualization they were actually looking for. Especially long tables, such as referrer or social share statistics take up a lot of space, but are used only by certain users.

Another issue, that was however not mentioned by the interviewees, is the inconsistent use of technologies. While a large part of the dashboard is realized as a client-side application being executed in the browser, there are also parts being rendered server-sided in the backend. As a result, the underlying code of these pages is spread across the frontend and backend code decreasing understandability and modularity of the corresponding codebase.

Functional and Technical Requirements

The overall goal of the dashboard is to prepare and visualize learner data in a way that supports teachers in decision-making, to increase their awareness of the students' learning process, to reflect on their teaching practice, to adjust it if necessary, and to enable problem detection. Thereby, it is important that the various visualizations are interpretable and can thus lead to action.

In terms of content, the general performance of the course has to be evident at first glance, i. e., the number of learners, 'no-shows', and, after completion, certificates. After that, it is important to visualize student activity and learning progress, i. e., what learning content is visited and how assignments are accomplished, to see if quizzes are too difficult or flawed. It is also important to gather feedback from students to be able to respond to problems. For this purpose, forum and data survey can be compiled. Also, geographic and demographic information are useful, to better understand the composition of the learning community.

From a technical perspective, the dashboard has to be performant, i. e., available data is displayed in an acceptable amount of time and more complex longer-running calculations do not block the entire rendering of the dashboard. Besides, the different aggregation levels of the data have to be reflected in the structure of the dashboard, i. e., detailed data visualizations and large tables are only displayed at subsequent pages and on request to not clutter the dashboard and overwhelm the teacher.

Revised Technical Concept

To meet the needs of the stakeholders and solve issues of the existing prototype, the entire concept of the dashboard is revised from different perspectives. In this paragraph, we focus on the technical requirements first, including the integration of WA metrics. The extension and completion of the dashboard with regard to the different functional domains of the platform is discussed in Section 4.3.

On a structural level the general goal is to clean up the existing dashboard to simplify the access to metrics and statistics. The actual objective of dashboards in general is to visualize complex data in a simple way to provide a quick overview about a certain topic, in this case the performance of specific courses. In the interviews it became clear that the existing dashboard contains a great number of different visualizations, whereas the majority of them is not relevant for all stakeholders. The new concept focuses on visualizations being relevant for the majority of the users while providing possibilities to obtain extensive information on demand. As KPIs are highly relevant for all interviewees, the corresponding part is retained. In contrast, detailed visualizations built for special purposes are moved to separate pages, referred as 'statistics pages' in the following. However, the information of the moved parts still have to be represented in the dashboard. Hence, there is a component in the dashboard visualizing the underlying data at a higher level for each statistics page, which takes up less space and is also easier to understand. At the same time, it serves as a link to the corresponding statistics page. For example, the list of social networks the courses have been shared in is moved to such a separate page. Along with this, the total number of course shares is added as a KPI to the dashboard. In this way, users receive an overview about the performance of a course and can follow the links in case they are interested in detailed information.

In addition to the revision of the existing concept, the new Google Analytics metrics introduced in Subsection 2.4.3 are integrated into the dashboard. This is achieved by adding new visualizations and creating new statistics pages, e. g., a heat map showing the average number of learners per hour per day of week. This component additionally links to a new activity statistics page. It shows histograms of session durations and the number of days between two subsequent sessions. The corresponding metrics group values to buckets, which ensures clarity and understandability of the visualizations. These bar charts also show the platform average for each bucket making it possible to compare the activity of a course with the average of all courses. Next to these two visualizations another heat map shows the number of active users for each day and hour in the course time frame. This visualization is similar to the heat map of the existing course dashboard that shows the temporal activity of users.

From a technical perspective, all pages are now rendered consistently on the server-side, but required analytics data is retrieved asynchronously through a dedicated API on the client-side. This approach has the advantage that the initial page is loaded quickly in the browser and the users already see the structure of the page while required data is loaded in the background. To retrieve all data that is visualized in the dashboard, multiple API requests to different endpoints are necessary. These requests are sent and processed concurrently. As a result, data is shown on the page as soon as it is received. Therefore, metrics which are calculated more quickly are already visible in the UI while more expensive operations are still running. In addition, all API requests are cached, each with an individual time frame based on the required timeliness of the data. Also, most of the data is lazy loaded, i. e., the request is delayed until the element is scrolled into the viewport and visible in the user's browser. All these improvements contribute to a better dashboard performance. In the backend, various queries of metrics are also optimized.

Listing 4.1: Example of a Data Visualization Pipeline

```
<ajax-wrapper data-url="/api/v2/statistics/course_dashboard/item_visits.json?
    course_id=265...dec" lazy-load="true">
  <data-selector key="count">
    <score-card icon-classes="eye" link="/courses/ruby2018/statistics/item_visits"
        link-text="More details" name="Item Visits"></score-card>
  </data-selector>
</ajax-wrapper>
```

To encapsulate and unify the retrieval, transformation, and visualization of the data in the frontend, a JavaScript library[1] is created to provide reusable Polymer[2] web components based on the popular libraries Plotly[3] and D3[4]. Thus, various custom HTML data source, transform, and receiver nodes can hierarchically pass the data to the respective subordinate nodes and thus form processing pipelines for the visualization of metrics. An example can be seen in Listing 4.1 which is rendered as shown in Figure 4.1.

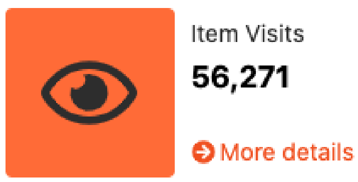


Figure 4.1: Rendered Data Visualization Pipeline for a Score Card Widget

[1] https://github.com/openHPI/m.e.i.n.e.l/
[2] https://www.polymer-project.org/
[3] https://plotly.com/javascript/
[4] https://d3js.org/

4.2.4 Evaluation

This subsection focuses on evaluating the technical enhancements of the dashboard. Therefore, the usability of the revised and extended LA dashboard is assessed by comparing it with the previous prototype on the basis of a conducted survey. This includes answering research question 3.2 by examining the usefulness of the implemented WA metrics.

Methodology

To achieve the goals of this evaluation, a survey was conducted addressing teaching team members of openHPI and openSAP as they are the target group of the according dashboard. As the audience is really specific and thus small, only 11 respondents could be acquired for answering the questionnaire. However, the importance of participants' views and positions are still highly relevant as they are experts in their fields, who utilize LA insights in their daily work. Therefore, results of this survey are still meaningful despite the small number of respondents. The participants were asked to express their agreement with the following ten statements separately for the existing and the revised dashboard:

Q01: The dashboard helps me to monitor the activity of my courses.
Q02: The dashboard facilitates access to relevant metrics.
Q03: The dashboard meets my needs.
Q04: I regularly use the dashboard for my work.
Q05: The dashboard is easy to use.
Q06: The dashboard is understandable.
Q07: The dashboard loads fast.
Q08: The dashboard is clear and tidy.
Q09: The dashboard works the way I would expect.
Q10: I like to use the dashboard.

A Likert scale with the following four levels and corresponding scores was utilized for giving answers: (0) strongly disagree, (1) somewhat disagree, (2) somewhat agree, and (3) strongly agree. The complete response data can be seen in Appendix C.4. For evaluating the significance of differences, a Wilcoxon test is performed based on the answers' scores for each question. Additionally, effect sizes are computed with Cohen's d. These statistics can be found in Table 4.1. The participants were also approached for qualitative feedback by means of free-text questions. This includes asking for suggestions for improvements. While the first two questions are considered in the following, the mentioned improvement suggestions are discussed later.

The survey results are analyzed concerning two aspects of this work. First, the usefulness of WA insights in context of LA is evaluated for answering research question 3.2. Afterward, the ease of use and satisfaction of the revised concept are assessed based on the answers of the participants to examine whether the goals of the revision are accomplished.

Usefulness of Web Analytics Insights

From a content-related perspective, integrating the implemented WA metrics did not result in a significant difference of the usefulness ($p < 0.05$). The new metrics were not highly relevant

Table 4.1: Descriptive and Inferential Statistics for Usability Before and After Revision of the Dashboard

	Existing Dashboard		Revised Dashboard		Wilcoxon	
Q	Mean	Std.Dev.	Mean	Std.Dev.	p-value	Cohen's d
01	2.1818	0.6030	2.2727	0.6467	0.6547	0.1454
02	2.0909	0.7006	2.4545	0.8202	0.1573	0.4767
03	1.6364	0.6742	1.9091	0.7006	0.1797	0.3967
04	2.4545	0.6876	2.5455	0.5222	0.5637	0.1489
05	1.9091	0.9439	2.3636	0.8090	0.1025	0.5170
06	1.3636	0.6742	2.1818	0.8739	0.0235	1.0484
07	1.0909	0.5394	2.0000	0.7746	0.0152	1.3621
08	1.2727	0.6467	2.3636	0.8090	0.0097	1.4895
09	1.8182	0.6030	2.1818	0.7508	0.0455	0.5340
10	1.6364	0.5045	2.2727	0.7862	0.0196	0.9633

for the majority of participants (Q02) and could neither help them in monitoring the activity in courses (Q01), nor increase the satisfaction of their needs (Q03) in a significant extent. Furthermore, the additional insights did not lead to respondents planing to use the dashboard more often in a remarkably scope (Q04).

The performance of the dashboard (Q07) was significantly increased ($p = 0.0152$) with a large effect ($d = 1.3621$). This was achieved partly by replacing certain existing metrics with appropriate WA metrics, which can be retrieved faster. Therefore, the integration of WA contributed to this improvement. However, the performance of the local analytics stores can also be boosted, e. g., by upgrading the underlying hardware. Especially for a non-commercial project, making use of the provided cloud infrastructure of WA tools, which is in case of Google Analytics even free of charge, is a meaningful decision to improve the performance of complex queries. The majority of the underlying metrics can also be covered by using WA methods. Consequently, WA can provide further insights and their usefulness is already proven. Nevertheless, the effect of these insights cannot be measured in this context as there is no basis of comparison. However, online learning platforms that do not have such sophisticated LA capabilities as the HPI MOOC Platform can benefit from integrating these WA metrics.

Ease of Use and Satisfaction of the Revised Concept

Besides examining the usefulness of the implemented WA metrics, the survey also aimed to ascertain the ease of use and satisfaction of the revised dashboard concept. In addition to the improved performance, the effect of which was already determined previously, the redesign led to a significant increase ($p = 0.0235$) of the understandability (Q06) with a proven large effect ($d = 1.0484$). Besides, a highly significant difference ($p = 0.0097$) was found in regard to the clearness (Q08) with a large practical effect as well ($d = 1.4895$). In terms of satisfaction of participants, a significant improvement ($p = 0.0455$) with an intermediate effect ($d = 0.5340$) was measured in the extent that the dashboard works as expected by the participants (Q09). Additionally, the respondents also prefer working with the revised version (Q10) as a significant difference ($p = 0.0196$) of respective answers with a large effect ($d = 0.9633$) was ascertained as well. However, the revision had no significant impact ($p = 0.1025$) on the simplicity (Q05).

In addition to the quantitative evaluation, the participants were asked to mention aspects of the revised concept they liked most. The majority brought up the improved clarity caused by moving detailed statistics to separate pages and enabling the possibility to drill-down. Besides, also the fact that data for each chart is loaded independently was well received. Other mentioned aspects were the availability of new KPIs and charts, the improved visualizations, and the increased number of LA insights.

4.2.5 Conclusion

In this section, it was shown which requirements for a learning analytics dashboard for teaching teams can be derived based on a prototype that has been developed in practice (research question 3.1). For this purpose, experts who had used the prototype extensively were interviewed and related work was examined. In addition to requirements based on the functional domains of the platform, technical issues were also uncovered. In particular, the performance and hierarchy of the visualized data are crucial. As a first step, a revised technical concept was therefore elaborated and implemented, which included especially the use of WA methods.

The evaluation of the usability of the revised and extended dashboard has shown that the newly implemented WA metrics did not have statistically significant differences in terms of the usefulness of the dashboards (research question 3.2). However, a large part of the existing LA capabilities, which are indeed proven to be useful for the stakeholders, can also be realized by using WA methods. Consequently, WA can still provide useful insights in the context of LA. In addition, it can contribute to an increase of the general performance by making use of the cloud infrastructure of WA tools. Besides, the revision of the dashboards had a significant impact on the ease of use and satisfaction. Especially the clearness was improved notably as shown by the quantitative, but also qualitative evaluation. Therefore, the intentions and goals of the revision were accomplished. The next iteration focuses on the implementation of the remaining functional requirements.

4.3 The Teacher Dashboard in Practice

Building on the technical improvements of the last study, the functional requirements for the teacher dashboard are completely implemented and examined in this section. In particular, the focus is on the usage during different teaching phases of a course (research question 3.3) and the perception of teachers with different experience levels (research question 3.4).

4.3.1 The Teacher Dashboard and Statistics

In this subsection, we present the different visualizations and statistics that are implemented for course teachers and managers on the platform to achieve the functional requirements described in Subsection 4.2.3. The provided visualizations and statistics are based on the findings of related work (Subsection 4.2.2), the technically available data in the platform, and the previously collected input and feedback from teachers. For this work, we explicitly limit the scope of the dashboard to the visualization of data. This means that the interpretation of data and the derivation of recommendations for action are up to the teacher. We retain the possibility of implementing such features in a future iteration of the system, which then allows to measure the impact of such extensions compared to this version.

Enrollments & Certificates Overview

The dashboard serves as a central entry point and presents the most relevant KPIs aggregated for the entire course. First, there is an overview of all enrollments and certificates (as soon as they are available) as shown in Figure 4.2. In addition to the current numbers, different points in time are shown for the course: the start, the end, and the middle. The middle of the course is a fixed date until which it is still possible to enroll in the course and achieve enough points to receive a Record of Achievement. Based on this value, the completion rate is calculated. In contrast, the Confirmation of Participation is determined by the number of students at the end of the course. Additionally, students are filtered for 'shows' (the opposite of 'no-shows')—i. e., users who have viewed at least one learning item. These figures are the usual KPIs used to communicate the success of a MOOC to the outside world and the higher management in organizations. Further details about these figures are discussed in Subsection 1.4.1.

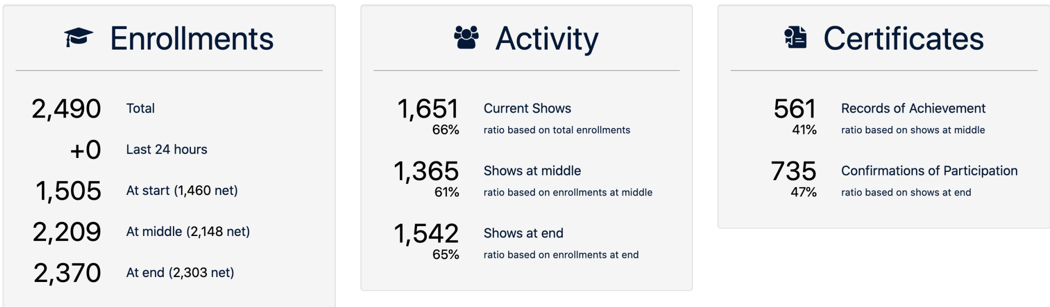


Figure 4.2: Dashboard: Enrollments and Certificates Overview for the Overall Course

Learning Items and Forum KPIs

Afterward, certain KPIs for the learning items and the forum are displayed, with which, e. g., the activity and difficulty of several MOOCs can be compared quickly. The learning items KPIs (Figure 4.3) show the number of total item visits, videos played by users, video asset downloads, link clicks in texts, graded quiz performance, and self-test performance. Each KPI provides a link

Figure 4.3: Dashboard: Learning Items KPIs for the Overall Course

to more details, which opens a statistics page with all individual items of the corresponding type, that shows further key figures per item. The forum KPIs (Figure 4.4) show the total number of posts and topics in the forum and how many learners were active in the forum.

Figure 4.4: Dashboard: Forum KPIs for the Overall Course

Enrollments & Learners and Forum & Helpdesk Historic Data

The historical development of some data is displayed as a graph. First, the growth of enrollments and learners is shown (Figure 4.5). In addition to the total enrollments, new learners on the platform (i. e., their first course) and 'no-shows', the enrollments of the last day, and the active users of the last day or the last 7 days can be displayed. The dotted vertical lines mark the start and end date of the course. Second, the historical growth of forum and helpdesk KPIs is shown in another graph (Figure 4.6), i. e., the total number of posts, topics, and helpdesk tickets.

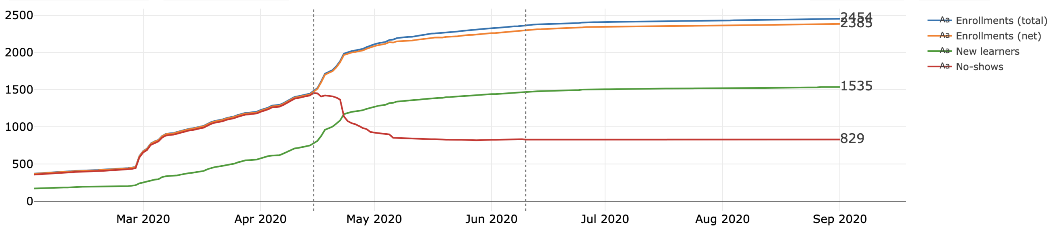


Figure 4.5: Dashboard: Enrollments and Learners Historical Data for the Course

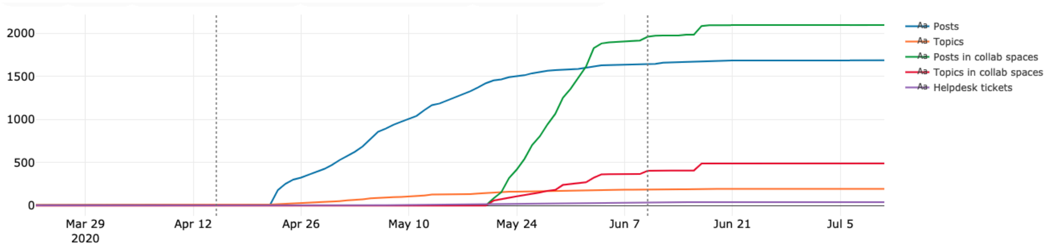


Figure 4.6: Dashboard: Forum and Helpdesk Historical Data for the Course

User Locations

The dashboard displays a world map with countries as shown in Figure 4.7. The coloring visualizes the number of users having accessed the course from the respective countries based on an IP address mapping. The user locations can, for example, help to identify and measure the reach of geographical target groups.

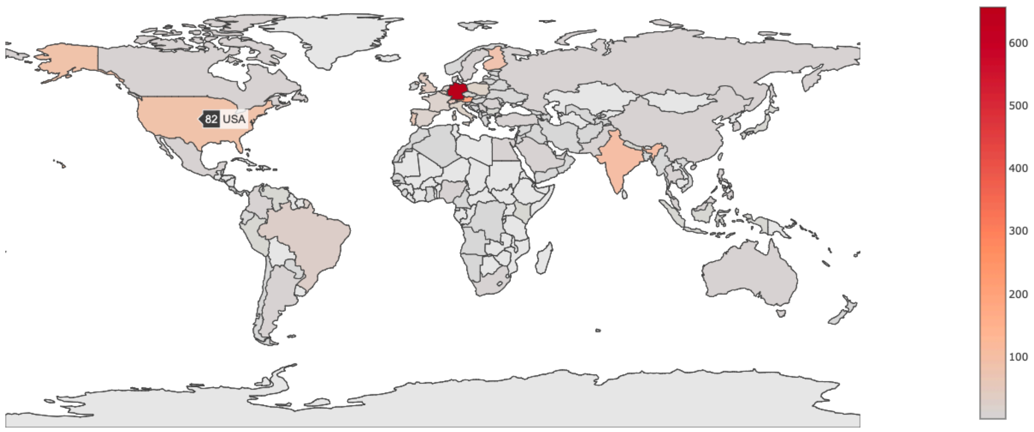


Figure 4.7: Dashboard: User Locations

Active Learners by Time of Day, Age Distribution, and Client Usage

At the bottom of the dashboard, there are three more visualizations (Figure 4.8). First, a heat map of active learners by time of day showing in aggregated form how many users were active on average per hour per weekday. Courses are usually structured in weeks, which often results in activity patterns on the platform, which can be traced here. Second, an age distribution is shown, which compares the different age groups of the course participants with the overall average of the platform. This makes it possible to examine whether the course addressed a rather younger or older target group. Third, a Venn diagram of the client usage is displayed, i. e., how many users accessed the course via a browser on a computer, a browser on a mobile device, or with one of the

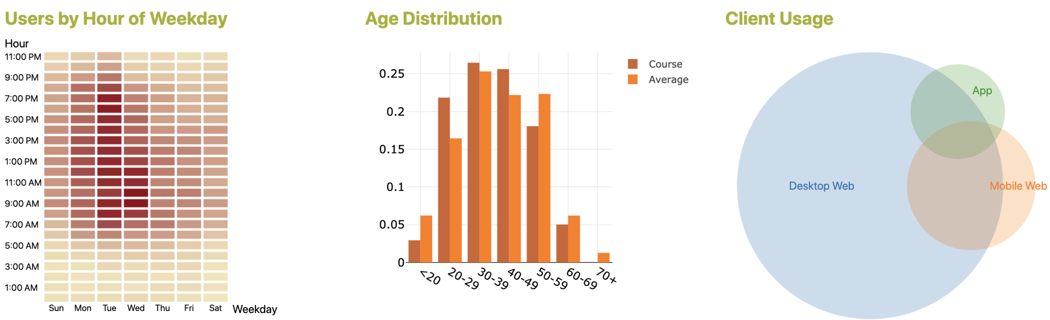


Figure 4.8: Dashboard: Active Learners by Time of Day, Age Distribution, and Client Usage

mobile apps for Android and iOS, as well as the intersections of users who used multiple devices. This enables to assess how many students have learned mobile or stationary.

In addition, the dashboard offers the possibility to open further statistics pages via a menu or links below the individual KPIs. These pages offer deeper insights into specific aspects of the course, mainly about the different learning contents and the students' interactions.

Item Visits

The item visits page offers a table with all learning items of the course with the number of visits and visits by unique users (Figure 4.9). A second table shows the average visits accumulated per item type. This page can be used to identify the most and least viewed content and content type.

Item	Position	Content Type	Exercise Type	Visits	Unique Users	Actions
0.1 Welcome to the Course	1.1	Rich text		3459	1513	More details
0.3 Module Overview	1.2	Video		1924	1138	More details
0.4 Course Elements	1.3	Rich text		1455	989	More details
0.5 Introducing the CORSHIP Team	1.4	Rich text		1975	985	More details
1.1 Welcome to Week 1	2.1	Rich text		4929	1346	More details
1.1 Welcome to Week 1	2.2	Video		2409	1204	More details
0.2 Pre-Course Survey	2.3	Quiz	Survey	3523	1238	More details
1.2 Corporate & Startup Collaboration: The New Normal	2.4	Video		2861	1194	More details
1.2 Corporate & Startup Collaboration: The New Normal	2.5	Quiz	Selftest	4966	1133	More details

Figure 4.9: Statistics Page: Item Visits

Videos, Video Downloads, and Video Details

The videos page provides a table with all video items of the course with the number of plays by users, the durations, the number of forward and backward seeks, and the average farthest watched percentage by users (Figure 4.10). The latter is based on the largest video timestamp per user captured during an interaction with the video player, e. g., by pressing play/pause or seeking in the video. This means that the user has not necessarily watched the entire video content up to this point, but it serves as an approximate value. This page can be used to detect videos that are often aborted or in which jumps are frequently made. The video downloads page displays for the same video items the number of downloads of the video streams, but also additional material like presentation slides or audio files (Figure 4.11). The video details page shows the presented KPIs for a single video and below the captured interactions with the video player (play, pause, seek, and changed speed) by users accumulated in chunks on a timeline (Figure 4.12). This can help to detect anomalies, e. g., indicating ambiguities in the video.

Item	Position	Played by Users	Duration	Average Farthest Watched by Users	Forward Seeks	Backward Seeks	Actions
1.2 Corporate & Startup Collaboration: The New Normal	2.4	973	03:48	100.00%	1233	678	More details Copy ID
1.1 Welcome to Week 1	2.2	935	01:43	98.63%	548	130	More details Copy ID
1.4 E-tivity: Myth-breaking	2.8	855	01:12	99.43%	246	109	More details Copy ID
1.5 Definition of Corporate & Startup Collaboration	2.10	811	05:58	100.00%	1750	2153	More details Copy ID

Figure 4.10: Statistics Page: All Videos of the Course

Item	Position	Total Downloads (by Unique Users)	HD Video Downloads (by Unique Users)	SD Video Downloads (by Unique Users)	Slides Downloads (by Unique Users)	Audio Downloads (by Unique Users)	Transcript Downloads (by Unique Users)	Actions
1.2 Corporate & Startup Collaboration: The New Normal	2.4	855 (361)	108 (90)	64 (49)	415 (308)	28 (24)	240 (188)	More details Copy ID
1.6 Expectations of Cooperation: Benefits and Potentials for Corporates	2.12	782 (346)	98 (76)	54 (43)	382 (290)	25 (20)	223 (177)	More details Copy ID
1.5 Definition of Corporate & Startup Collaboration	2.10	741 (343)	92 (74)	48 (39)	368 (294)	22 (19)	211 (178)	More details Copy ID
1.6 Expectations of Cooperation: Benefits and Potentials for Startups	2.13	680 (313)	83 (65)	50 (40)	336 (268)	18 (17)	193 (166)	More details Copy ID

Figure 4.11: Statistics Page: All Video Downloads of the Course

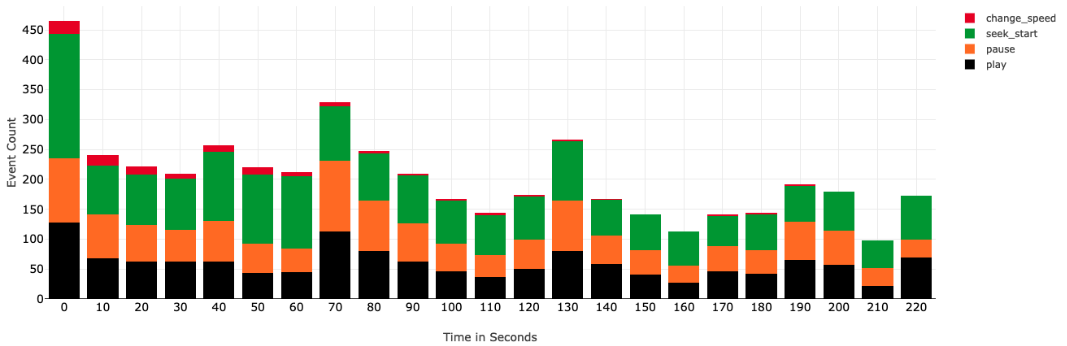


Figure 4.12: Statistics Page: Details of a Video (Interactions)

Quizzes and Quiz Details

The quizzes page provides a table for all graded quizzes (Figure 4.13), self-tests, and surveys of the course. For each quiz, the overall submission count, the average performance, and the average submit duration are presented to quickly identify problematic items. As for all these tables, it can be easily navigated to the quiz details page of a single quiz. Here, a historical graph is shown with all submissions over time (Figure 4.14), as well as statistics for each question of a quiz (Figure 4.15). These usually provide the number of correct and wrong submissions of a question, the average points, and a bar chart with the number of all given correct and wrong answer options. All of this can be used to determine if a quiz matched the expected level of difficulty, or if a question or answer was too challenging or ambiguous and should be re-graded.

Graded Quizzes

Item	Position	Submission Count	Unique Users	Average Performance	Average Submit Duration	Actions
1.8 Weekly assignment	2.17	744	744	81.72%	790 s	More details
2.11 Weekly Assignment	3.22	652	651	69.96%	996 s	More details
3.22 Weekly Assignment	4.24	625	622	84.16%	824 s	More details
4.15 Weekly assignment	5.21	609	608	76.08%	945 s	More details
Final Exam	6.1	610	607	79.20%	2465 s	More details

Figure 4.13: Statistics Page: All Graded Quizzes of the Course

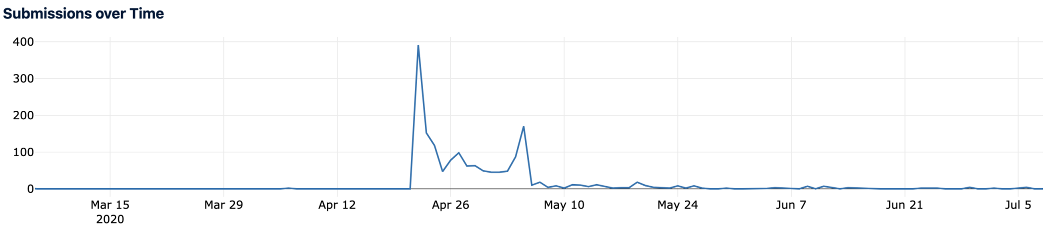


Figure 4.14: Statistics Page: Details of a Quiz (Submissions over Time)

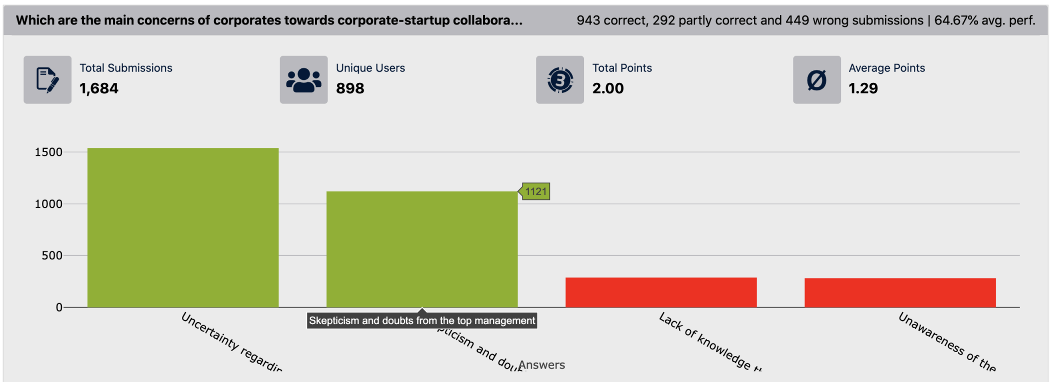


Figure 4.15: Statistics Page: Details of a Quiz (Answers and Performance of a Question)

Texts and Text Details

The possible interactions with text items are very limited, so the texts and text details page can only provide information about the number of clicked hyperlinks by users, if there are any (Figure 4.16).

Item	Position	Total Clicks	By Unique Users	Earliest Click	Latest Click	Actions
0.1 Welcome to the Course	1.1	346	211	Feb 19, 2020 1:32 PM	Nov 27, 2020 4:52 PM	More details Copy ID
5.2 Team Assignment Description	7.3	494	121	Feb 14, 2020 1:00 PM	Aug 17, 2020 10:53 AM	More details Copy ID
1.1 Welcome to Week 1	2.1	125	102	Dec 20, 2019 12:51 PM	Dec 8, 2020 7:10 PM	More details Copy ID

Figure 4.16: Statistics Page: All Texts of the Course

Activity

The last page that is part of this work is the activity page (Figure 4.17). Similar to the active learners by time of day on the dashboard, here the active learners per hour are displayed in a heat map, but for each calendar day during the full course runtime, instead of aggregated per weekday. Therefore (missing expected) activity peaks for every day of the course can be identified with this visualization.

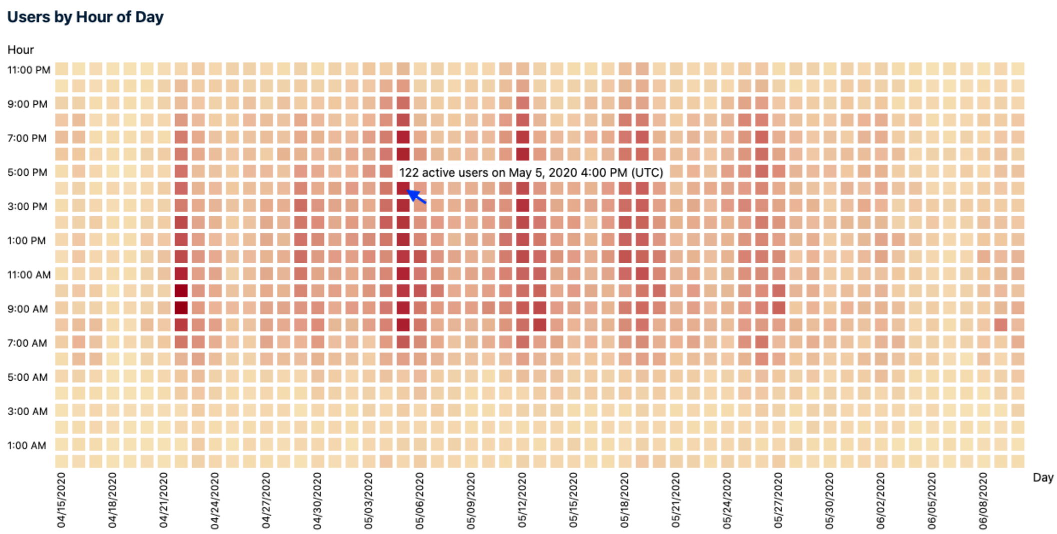


Figure 4.17: Statistics Page: Activity (Users by Hour of Day)

4.3.2 Evaluation

In the following, we present an analysis of the usage data of the teacher dashboard and statistics, and the perception based on survey results to examine the two remaining research questions 3.3 and 3.4. By means of the findings, we also suggest future improvements and research directions.

Methodology

To investigate the research questions, data were collected in two ways. First, automatically captured usage data were collected from teachers and managers within the platform when they visited the dashboard and statistics pages. This was conducted in the period from January 1, 2020, to October 1, 2020. All publicly and freely accessible courses that ran on three selected instances of the HPI MOOC Platform, which both started and ended during this period, are used for analysis. Thus, the usage data of 33 courses and 4,461 dashboard and statistics page visits is evaluated. The entire data set is available in Appendix C.5.

Second, these teachers and managers were invited to a voluntary survey, which 23 of them completed. It turned out to be difficult to get this group of people to answer an extensive survey, as none of them teach or manage MOOCs full-time. Often, it is only a small part of their everyday work. The first part of the survey addressed the evaluation of the usefulness of the various visualizations and statistics during different teaching phases of a course. For this purpose, we use the classification of the MIT Online Course Design Guide [90], which is structured into four phases and consists of:

Pre-Design:

- The readiness of existing course content
- The experience with online teaching and learning
- Learner target analysis

Design and Development:

- Objectives and outcomes
- Course structure
- Lecture development
- Assessment design

Facilitation:

- Course announcements and guides
- Forum administration
- Community engagement
- Troubleshooting and content updates

Evaluation:

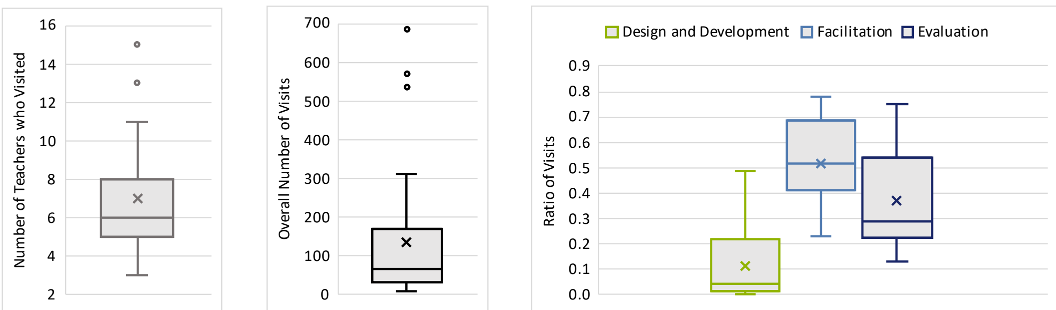
- Course feedback
- Assessing analytics and statistics

We use the same classification for the analysis of the usage data, but only for the last three phases since there is no data available during the pre-design phase. In this phase, usually no work is being done on the platform and therefore there is no course or page to visit yet. With the second part of the survey, we examine the overall perception of the dashboard and statistics based on the Evaluation Framework for Learning Analytics by Scheffel [131], which provides quality indicators for LA applications. We also subdivide and compare the results according to the different teaching experience levels in MOOCs of the participants, which were also questioned in the survey. The complete survey responses can be found in Appendix C.6.

Usage of the Teacher Dashboard and Statistics

Based on the usage data of nine months, it can be seen in Figure 4.18a that an average of 7 teachers and managers per course used the dashboard and statistics pages, with a minimum of 3 and a maximum of 15 people. The pages were visited between 8 and 687 times, with an average of 135 hits per course (Figure 4.18b). In about 11.5% of the cases, the accesses took place before the courses started, in the design and development phase; in about 51.5% of the cases during the courses' runtime, in the facilitation phase; and in about 37% of the cases after the courses ended, in the evaluation phase (Figure 4.18c).

Research question 3.3 can be answered to the effect that the dashboard and statistics pages are used very frequently and by several teachers and managers per course, mainly while a course is running and afterward for evaluation, but rather less during the design and development of a course. This is plausible since no data is available before the course starts. However, data from other courses or previous iterations of the same course can be used to make design decisions, e. g., regarding the expected target group or to improve problematic content. The use during the evaluation phase is given. Nevertheless, these correlations between courses cannot be automatically captured yet in the platform and are therefore not part of this analysis. Overall, the high usage numbers show that the dashboard and statistics pages are accepted by teachers and managers and are an integral part of their processes when working with the platform.



(a) Number of Teachers Who Visited the Pages per Course

(b) Overall Number of Page Visits per Course

(c) The Ratio of the Teacher Dashboard and Statistic Page Visits in Three Different Teaching Phases per Course

Figure 4.18: Variations in the Number of Visits of the Teacher Dashboard and Statistics Pages in the Examined Courses ($N = 33$)

Perception of the Teacher Dashboard and Statistics

Regarding the demographics of the 23 survey participants, about 65% were male and 35% female. In terms of age distribution, 22% said they were between 20-29 years old, 57% between 30–39, 17% between 40-49, and 4% between 50–59. Furthermore, 39% declared that teaching, in general, is a regular part of their work. So we are mainly faced with experienced professionals who have very different expertise in teaching. In the first part of the survey, we asked about the usefulness of all visualizations and statistics, as presented in Subsection 4.3.1, according to the following scheme: "How useful do you find the <visualization> in the different course phases?" The four course phases pre-design, design and development, facilitation, and evaluation were introduced and explained to the participants beforehand. The answers were captured with Likert scales with a numerical range from 1 (not useful at all) to 7 (very useful). The average value for each dashboard visualization is shown in Table 4.2 and for each statistics page in Table 4.3. Additionally, we illustrate the values in Figure 4.19 as radar charts.

Table 4.2: Average Usefulness Scores of the Teacher Dashboard Visualizations in the Four Teaching Phases

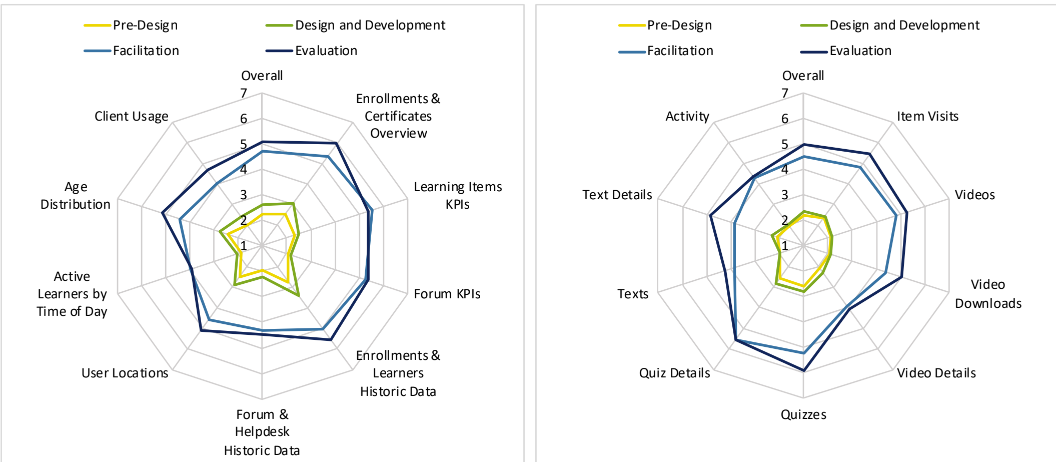
Phase	V1	V2	V3	V4	V5	V6	V7	V8	V9	Overall
Pre-Design	2.57	2.35	2.04	2.74	1.96	2.48	1.87	2.43	2.00	2.27
Design and Development	3.09	2.48	2.17	3.39	2.22	2.87	2.04	2.78	2.43	2.61
Facilitation	5.35	5.52	5.26	5.04	4.30	4.57	3.96	4.43	4.00	4.71
Evaluation	5.96	5.39	5.35	5.57	4.48	5.13	3.91	5.13	4.65	5.06

V1: Enrollments & Certificates Overview. V2: Learning Items KPIs. V3: Forum KPIs.
V4: Enrollments & Learners Historic Data. V5: Forum & Helpdesk Historic Data. V6: User Locations.
V7: Active Learners by Time of Day. V8: Age Distribution. V9: Client Usage.

Table 4.3: Average Usefulness Scores of the Statistics Pages in the Four Teaching Phases

Phase	S1	S2	S3	S4	S5	S6	S7	S8	S9	Overall
Pre-Design	2.35	2.13	2.09	2.09	2.61	2.57	1.96	2.09	1.96	2.20
Design and Development	2.43	2.17	2.13	2.30	2.83	2.87	1.96	2.30	2.04	2.34
Facilitation	4.78	4.83	4.39	3.91	5.22	5.57	3.83	3.83	4.26	4.51
Evaluation	5.43	5.26	5.04	4.09	5.91	5.57	4.22	4.83	4.35	4.97

S1: Item Visits. S2: Videos. S3: Video Downloads. S4: Video Details. S5: Quizzes. S6: Quiz Details. S7: Texts.
S8: Text Details. S9: Activity.



(a) Dashboard Visualizations

(b) Statistic Pages

Figure 4.19: Perceived Average Usefulness of the Teacher Dashboard Visualizations and Statistic Pages in the Four Teaching Phases

For the dashboard visualizations (Figure 4.19a), it can be seen that the items regarding enrollments and learners, certificates, geographical data (user locations), and demographical information (age distribution) are considered most useful. For the statistics pages (Figure 4.19b), especially the quiz pages are considered most useful, followed by the item visits and video pages. All in all, it is visible that the usefulness is rather limited in the pre-design (2.27 and 2.20 overall) and design and development phase (2.61 and 2.34 overall), whereas a higher usefulness is perceived during the facilitation (4.71 and 4.51 overall) and evaluation phase (5.06 and 4.97 overall). The results are in line with the usage statistics, which allows the same interpretations. Before a course has started, the lack of data and comparisons with already completed courses reduces the usefulness of the dashboard and statistics pages. Some participants only thought of this possibility during the survey, as they noted in the open questions that followed. However, once data on the actual course are available, its usefulness increases. How comprehensible and beneficial these insights are, is questioned in the next part of the survey.

Based on the EFLA questionnaire [131], we surveyed eight quality indicators for learning analytics tools grouped by the three dimensions: data, awareness and reflection, and impact. The answer options were based on a Likert scale with numerical values from 1 (strongly disagree) to 10 (strongly agree). As the framework proposes, we calculate the average values for each question, then the dimensional scores by rounding the result of $((x-1)/9)*100$ where x is the average value of a dimension, and finally the overall EFLA score as the mean of the three dimensional scores. The questions were:

Data:
- D1: For the Course Dashboard and Statistics it is clear what data is being collected.
- D2: For the Course Dashboard and Statistics it is clear why the data is being collected.

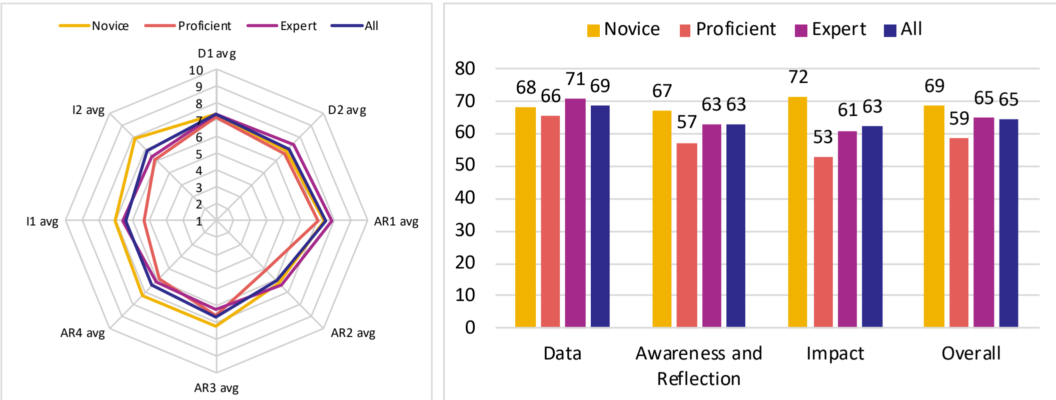
Awareness and Reflection:
- AR1: The Course Dashboard and Statistics make me aware of my students' current learning situation.
- AR2: The Course Dashboard and Statistics make me forecast my students' possible future learning situation given their (un)changed behavior.
- AR3: The Course Dashboard and Statistics stimulate me to reflect on my past teaching behavior.
- AR4: The Course Dashboard and Statistics stimulate me to adapt my teaching behavior if necessary.

Impact:
- I1: The Course Dashboard and Statistics stimulate me to teach more efficiently.
- I2: The Course Dashboard and Statistics stimulate me to teach more effectively.

We suspect that the interpretation and derivation of actions from the dashboard and statistics pages depend on the experience level of the participants. To investigate this in more detail, we also divide the results according to the question: "What is your experience as part of a teaching team in MOOCs?" with following answer options:

Novice: I have only developed and taught 1 MOOC so far.
Proficient: I have developed and taught 2 to 5 MOOCs so far.
Expert: I have developed and taught more than 5 MOOCs.

The participants showed a balanced distribution of teaching experience levels in MOOCs: according to our classification, 35% were novices, 26% were proficient, and 39% were experts. The results of the different quality indicator mean values and dimensional scores are displayed in Figure 4.20. Among the novices, the reason for data collection (D2: 6.88) and forecasting the students' learning situation (AR2: 6.25) received the lowest ratings, whereas increasing the awareness of the students' learning situation (AR1: 7.38) and the effectiveness of teaching (I2: 7.88) received the highest ratings. Among the proficient, also forecasting the students' learning situation (AR2: 5.17) and increasing the efficiency of teaching (I1: 5.33) were rated worst, and increasing the awareness of the students' learning situation (AR1: 7.00) and the type of data collection (D1: 7.17) best. Among the experts, the adaptation (AR4: 6.11) and reflection of teaching (AR3: 6.22) received the lowest ratings, whereas the reason for data collection (D2: 7.44) and increasing the awareness of the students' learning situation (AR1: 7.89) received the highest ratings.



(a) Average Values of the EFLA Items

(b) Average Scores of the EFLA Dimensions

Figure 4.20: EFLA Results for the Teacher Dashboard and Statistic Pages for the Different Teaching Experience Levels in MOOCs

Altogether, the aspect of forecasting the students' learning situation (AR2: 6.04) is improvable in particular, whereas increasing the awareness of the students' learning situation (AR1: 7.48) represents the greatest benefit of the dashboard and statistics pages for teachers and managers. However, the values are not too divergent overall and in general within a neutral to a moderately positive area. Almost no aspect was rated negatively, which is also indicated by the dimensional scores in Figure 4.20b (100 is the highest possible value, 0 the lowest). There are no practically relevant differences between the dimensions, but a tendency can be seen in all dimensions between the different experience groups. Novices have the highest scores (overall 69), proficients the lowest (overall 59), and experts are in between of them (overall 65). This leads to the interpretation that teachers with less experience explore the possibilities with a certain curiosity. Proficients become temporarily a bit disillusioned, and experts, who know better what is useful for them and what is not, become more productive again in working with the provided tools. There is one exception in the data dimension: here, experts have a higher score (71) than novices (68). Since this is all about the "what" and "why" of data collection, it seems that with increasing experience, the understanding of relationships between collected data and the available visualizations and statistics in the platform is increasing. In total, the overall EFLA score is 65. This enables fellow researchers and us to compare similar tools with our dashboard and statistics pages and to measure improvements in future iterations.

Research question 3.4 can be answered to the effect that teachers and managers perceive the dashboard and statistics pages most useful during the facilitation and evaluation of a course, but less before, which matches their usage behavior. The greatest benefit of these data-driven insights is the increased awareness of the students' learning situation. The level of experience in teaching has only a minor impact on the perception, although teachers with little or very much practice in MOOCs find these tools slightly more helpful for their work than those with moderate experience.

Future Work

The results showed that the dashboard and statistics pages still have potential during the design and development of a course. On the one hand, training can be offered for teachers to explain how the dashboards from previous courses can be used to improve the design of their courses. On the other hand, since this cannot be offered in a scalable way for everyone because of the sheer amount of MOOCs, technical support is needed. This requires the dashboard to include comparison features with other courses, which also has the advantage that the usage in the phases before a course starts can be measured and analyzed more effectively.

In order to better support not only awareness, but also reflection and adaptation of the teaching behavior, the next iteration of the dashboard has to offer more explanations of the presented metrics and recommendations for action. Furthermore, automated predictions can be provided to better forecast the future learning situation of students and to react to issues early on to make teaching in MOOCs even more efficient and effective.

4.3.3 Conclusion

In this section, we presented a subsequent study of a learning analytics dashboard for teachers in MOOCs as these tools are vital in the development, facilitation, and evaluation of courses. The dashboard and subordinate statistics pages were implemented for the HPI MOOC Platform based on an examination of related work, previously collected feedback and input from teachers working with the platform, and the technically available data. Visualizations were developed regarding course enrollments and active students, student performance in assessments, use of learning materials like videos, forum activity, demographic and geographic information, and activity patterns.

The mixed-method evaluation was divided into two parts: platform data were utilized to analyze the usage of the dashboard and statistics pages (research question 3.3), and self-reported survey data were used to assess teachers' perception of these tools (research question 3.4). Based on the data of 33 courses, we discovered that the dashboard and statistics pages are used very frequently (135 times on average per course) and by several teachers and managers (7 users on average per course). The accesses occur mainly during the course runtime and especially afterward, as part of the course evaluation. The high usage figures emphasize the acceptance of the dashboard and statistics pages by teachers and are considered as an integral part of their teaching processes when working with the platform. However, they are rarely used during the design and development of courses since there is almost no data available at this phase. Insights from previous courses can be utilized, but this usage cannot be technically captured and linked yet.

Based on survey responses, we found that the perception matches the results from the usage analysis: teachers and managers consider the dashboard and statistics pages most useful during the facilitation and evaluation of a course, although less useful before. It helps them to increase their awareness of the students' learning situation, whereby the support to reflect and adapt their teaching behavior can be improved. The level of experience in teaching has little effect on the perception, even though teachers with limited or extensive experience in MOOCs perceive these tools more beneficial than those with moderate experience. In the future, a solution can be implemented allowing teachers to make comparisons with other courses within the dashboard and to provide better support during the design and development of a course. Also, recommendations and more explanations of the visualizations can help to make the data-driven insights more actionable for teachers.

All in all, we are confident that this study provides a valuable contribution to the still young research field of learning analytics dashboards for teachers in MOOCs. We presented further evidence of the benefits and need for data-driven insights to teach at scale, the design and evaluation of such tools, and further challenges.

4.4 Summary

Based on the learning analytics capabilities of the HPI MOOC Platform, this chapter examined how data-driven insights can also support teaching in MOOCs. The focus was set on technology-enhanced monitoring of the students' learning as it is a key method in teaching. However, personal supervision is barely feasible in MOOCs. For this purpose, a learning analytics dashboard for teaching teams was developed in multiple iterations and improved based on user feedback. Thereby, four aspects were formalized as research questions and investigated.

An initial prototype of the dashboard emerged from the daily use of the platform and the repeated data requests from various stakeholders. These were gradually implemented in a dashboard and related mainly to high-level KPIs such as the number of learners in a course. Nevertheless, an overall concept was lacking, which is why a systematic review of issues and the definition of requirements were identified as the first aspect. For this purpose, the approach was twofold. First, related work from research and practice about dashboards for teachers in MOOCs was examined and its findings analyzed. Second, expert interviews were conducted with personnel from openSAP who work primarily with the platform and develop and host courses. Based on these insights, various functional and technical requirements were derived. The overarching purpose of the dashboard is to prepare learner data in a way that supports teachers in decision-making, increasing their awareness of the learning progress, stimulating reflection on their own teaching, and if needed providing impetus when problems require adjustments and action. Regarding content, different functional domains of the platform need to be addressed. In addition to overall learner numbers, 'no-shows', and certificates, insights into activity and learning progress are required, i. e., how learners visit the content and perform in assignments. It is also important to provide indicators for and be able to respond to feedback and questions in the forum and surveys. Lastly, geographic and demographic information are also useful to better understand the composition of all students. On the technical side, there was criticism that the dashboard needs to be more performant. Up to this point, the dashboard was only displayed as a whole when all data was processed and available. It is inconvenient that certain longer running calculations block the visualization of immediately available data. In addition, it was noted that the previous prototype is too cluttered. Therefore, the presentation of more detailed data has to be moved to subordinate pages and only be accessed when needed.

In a first improvement step, the technical requirements were implemented. Here, metrics based on the already introduced integration of web analytics methods were also used experimentally. Whether these can increase the usefulness of the dashboard for teachers was examined as a second research aspect. For this purpose, a survey was conducted with people from teaching teams of the openHPI and openSAP platforms, comparing the initial prototype and the revised dashboard. No statistically significant differences in terms of usefulness were found for the additional use of WA methods, nor was any increased practical relevance reported as a result of their integration. However, they were able to contribute to the perceived significant increase of the performance of the dashboard with a large effect size. Nevertheless, it has to be be mentioned that other technical solutions can also enable this, but WA platforms can outsource the required resources which saves costs and effort. In addition, significant improvements in understandability, clearness, and satisfaction were measured. The new structure, where more detailed data is moved to subordinate pages and linked from the dashboard, was also well received. Overall, the results showed that the intended technical improvement of the dashboard was achieved.

In the next iteration, the functional requirements were completely implemented and a detailed overview of all visualizations of the dashboard and statistics pages was presented. These include an enrollments and certificates overview, KPIs about the learning items and the forum, historic charts of the enrollments, learners, forum and helpdesk, a world map of the students' locations, a heatmap of active learners by the hour of weekday, a bar chart of the students' age distribution, and a Venn diagram of the used mobile and stationary client applications. The statistics pages include more detailed information about the students' visits of the learning items, interactions and downloads of the video lectures, statistics of quiz submissions, interactions with text items, and active users by the hour of day across the timeline of the course. As the third and fourth research aspect, we examined how teachers at different phases of a course use the dashboard and how they perceive it depending on their experience level. For the first study, the dashboard's usage data of teachers were analyzed. The dashboard and statistics pages were used very frequently and by multiple teachers per course, but mainly during its runtime and afterward, less during the design and development of it. At this point, there are almost no data available in the dashboard. It was noted that the usefulness can be increased in this phase by having comparison possibilities with other courses. Nevertheless, the usage figures show that it already is an integral tool for teaching in MOOCs. Afterward, the perception was examined by means of a survey. First, we analyzed which visualization is most useful in each teaching phase. The results are consistent with those of the previous usage evaluation. The usefulness generally increases with the availability of more data after a course has started. In detail, for the dashboard, the enrollment, learner and certificate figures, and the geographical and demographical data are found to be the most useful. For the statistics pages, the quiz statistics are perceived most useful, followed by the item visits and video figures. Second, we questioned eight quality indicators based on the Evaluation Framework for Learning Analytics. Thereby, the participants were classified according to their teaching experience with MOOCs as novice, proficient, or expert. The greatest benefit of the dashboard is the awareness of the students' learning situation, while the forecasting of the students' learning situation has the greatest potential to be improved. The experience level has only little impact on the perception. The main shortcomings of the dashboard are the possibility of comparison with other courses and recommendations for action, which have to be addressed in future iterations.

Overall, we provided further evidence of the benefits and need for learning analytics dashboards as one approach to support teaching in MOOCs with data-driven insights. Furthermore, we analyzed how such a tool is used and applied in the day-to-day teaching of a real-world MOOC platform, and which challenges and requirements arose during the development process.

5 General Discussion

The research focus of this dissertation was to investigate the technical integration and application of learning analytics in MOOCs to support learners and teachers with data-driven insights. For this purpose, various research prototypes were iteratively developed and examined in case studies. Thereby, the HPI MOOC Platform served as the learning environment, allowing the prototypes to be tested with users in authentic real-world learning and teaching situations.

5.1 Main Findings

The studies and experiments conducted in the course of this thesis served to answer the main research questions formulated in the general introduction. The findings acquired in the process are summarized in the following according to said questions.

Research Question 1: How can learning analytics be enabled at the scale of MOOCs?

To answer this question on the example of the HPI MOOC Platform, a generic learning analytics infrastructure was integrated into the platform's service-oriented architecture, which enables the provision of LA capabilities for all stakeholders of MOOCs. Building on the experiences and lessons learned in over five years of delivering LA for learners, teachers, researchers, and managers on multiple platform instances, we have compiled a set of design recommendations for platform vendors and researchers. These best practices support their decision-making when implementing LA features in MOOC platforms.

- Concurrent Data Collection and Processing: The performance impact on the overall application caused by additional analytics tasks has to be kept to a minimum. A common technique is to execute such tasks concurrently. This is realized by utilizing an asynchronous message queue for event collection to avoid blocking the sending components. The data processing is handled by a separate service running independently from other system components.
- Schema-Agnostic Pipelining: Different data schemas and query requirements fit more or less well to different storage technologies. Therefore, various analytics data are stored eventually in multiple databases. Hence, we propose a pipeline processing architecture. By utilizing an ETL process for this, all data can be processed based on a generic data schema. This enables a schema-agnostic data processing and minimizes technology and vendor lock-ins.
- Reusable Pipeline Components: By utilizing the proposed schema-agnostic pipeline architecture, all transformation processing steps become reusable. For example, this allows applying the same anonymization step to all analytics pipelines. This reduces implementation and maintenance efforts by applying the *don't repeat yourself* principle.

- Central Interface for Data-Driven Insights: Instead of having each application component providing its own analytics interface, it is reasonable to have a central interface for data-driven insights. This is realized with an index of all available metrics within the LA service. Also, it abstracts the underlying database technology.
- Embrace Open Standards: Interoperability with other applications and systems can be achieved best through the use of open standards. In the domain of LA, the xAPI format is accepted widely. This standard also defines the learning record store. Thus, an implementation of such an analytics store can be used right away without further data transformations.
- Data Protection by Design: By taking data protection into account in every project stage, privacy risks are reduced and trust increased. Users must stay in control of their data and the benefits of capturing and processing personal data have to be communicated beforehand. It must also be ensured at an early stage that legal requirements like GDPR are complied with.

In addition to the web-based access of the HPI MOOC Platform, users can also learn with mobile devices. For this purpose, native mobile applications for Android and iOS are available. Especially for lifelong learners, context is a key factor to integrate learning sessions in their daily life with a wide range of private and work activities. To better understand and support mobile learning, we developed mobile learning analytics capabilities that take the learners' context and requirements of mobile applications into account. Based on this, we proposed two architectural enhancements.

- Enrich Context Data: By integrating a context model into all captured analytics events, it is possible to identify different learning situations and adapt the learning experience accordingly. For this purpose, we enrich information about the used device and application, local time and place, and type of the connected network.
- Mobile Data Collection: Learning with mobile devices has some specifics which must be considered. In particular, offline learning has to be supported by client-side data persistence, and mobile data usage has to be reduced by batch transfers and prioritization of transmissions based on the network type.

Lastly, the LA infrastructure was extended by an external web analytics provider, as web analytics already provide sophisticated methods and are used productively in many industries to support data-driven decision-making. Although the foundation of both fields is similar, WA was not profoundly used for LA purposes so far. Based on a server-side proof of concept integration of Google Analytics, we investigated potentials and boundaries based on conceptual differences between both fields, while still maintaining data privacy.

- Limited Use of Web Analytics: Several aspects of LA can also be accomplished by utilizing WA. Nevertheless, the applicability of WA strongly depends on the type of stakeholder it is intended for. Especially highly aggregated metrics, such as information about the general activity and progress of a course, can easily be obtained with WA. However, limitations are reached when attempting to retrieve metrics about smaller groups or individual learners. Therefore, WA can be utilized mainly for platform owners, managers, and partially for teachers and researchers, but almost no use case can be provided for learners themselves.

Research Question 2: How can data-driven insights support learning in MOOCs?

To address this question, we provided technical support for self-regulated learning through data-driven insights, as this metacognitive skill set was identified as an important factor positively associated with students' achievement in online learning environments. Thereby, the focus was on the strategies of goal setting, strategic planning, and self-evaluation. For the latter two, a learning analytics dashboard for learners was developed to replace the existing progress page of the platform. The initial design was based on findings of related work, identified platform requirements, and input from an ideation session with experts. To study the learners' perceived usefulness and acceptance, and which visualizations are valued the most, we conducted a survey in a course which was completed by 217 learners.

- The learning analytics dashboard is considered as extremely useful by 84.79% of the learners.
- The most valued components are the progress overview (1.54 on a scale from -2 to +2), the course dates (1.33), the achieved points over time (1.31), the quiz performance (1.30), and repetition suggestions (1.14).
- Learners need detailed and easy-to-understand explanations of the displayed visualizations, including recommendations for action, and they want to customize the amount of displayed data.

Based on the initial findings, the dashboard was revised. The components were rearranged within their functional domain according to their perceived value. Also, further explanations were added and one component was entirely removed. Afterward, an A/B/n test was performed in five courses with three experiment groups: (1) learners that are able to use the old progress page ($N = 3{,}448$), (2) learners that are able to use the new progress overview but not the additional dashboard widgets ($N = 3{,}440$), and (3) learners that are able to use the complete revised learner dashboard ($N = 3{,}433$). We compared the achieved points and visited learning items between the groups to study completion rates and examined statistically significant differences utilizing a one-way ANOVA test. Also, group 3 was invited to a voluntary survey to evaluate whether the dashboard supports learners in applying self-regulated learning. It was based on the Evaluation Framework for Learning Analytics [131] and completed by 296 learners.

- No statistically significant differences are found in all five courses for both metrics, the visited items and achieved points.
- No statistically significant differences are found for a repeated analysis of different user cohorts within the same test groups: first-enrolled users, users with an easy to intermediate regular computer use, users older than 50, and users older than 60. We assumed that said user cohorts need more guidance in completing online courses.
- Hence, there are no differences in the completion rates of learners with regard to the use of the three dashboard variants. All three progress and dashboard variants are sufficient to achieve a certificate-based learning outcome.
- The dashboard's support for self-regulated learning is achieved especially for the strategy of self-evaluation by the stimulation of awareness (EFLA-AR1: 8.15) and reflection (EFLA-AR3: 7.11) of the learning situation and behavior. Besides, strategic planning is partially encouraged by stimulating the adaptation of the learning behavior when necessary (EFLA-AR4: 6.83).

To elaborate a concept for the strategy of goal setting, the next study examined how successfully learners achieve their self-reported learning objectives. For this purpose, related work was examined and five courses of the HPI MOOC Platform were analyzed. Here, learners could select their learning objective in a pre-course survey: (1) I would like to receive a Record of Achievement in the end and learn the course content; (2) I am mainly interested in learning the course content and the Record of Achievement is not important to me; (3) I am only interested in selected learning units; and (4) I just want to look around. The achievement of the learning objectives was determined by means of an evaluation of platform data. A total amount of 9,698 users provided their learning objective through the pre-course survey resulting in a response rate of 69.95%.

- 26.63% of the users are interested in a graded certificate and considerably 61.54% of the users are mainly interested in the content itself without the need of a certificate.
- 11.82% of the users are only interested in selected learning units or only want to look around.
- 49.90% of the users achieved or exceeded their goals and 50.10% missed their objective.
- Hence, a large user group either changes their goal during course runtime or drop out.
- The varying individual achievement rates across the different courses and found in literature point to the fact that goal achievement strongly depends on the course design, examination, and difficulty of different goals.

Based on the found capabilities and shortcomings of goal setting and achievement in MOOCs, a concept for personalized learning objectives was proposed with a technical focus on feasibility and automation, which was then implemented for the HPI MOOC Platform. In an initial evaluation that incorporated features for objective selection and guidance, the learners' acceptance was examined with an A/B/n test in two courses. Four experiment groups were defined: (1) this control group was not able to select a learning objective ($N = 2{,}293$); (2) this group saw an objective selection modal directly after enrolling in a course ($N = 1{,}126$); (3) this group was automatically prompted with the objective selection modal when visiting the learning content the first time and an infobox was added at the top of each item page, to open the modal again in case it was dismissed ($N = 2{,}399$); and (4) this group only saw the infobox at the top of each item page and therefore explicitly had to click on the provided link to see the objective selection modal ($N = 2{,}397$). After the selection of an objective, the corresponding learning content was highlighted in the course navigation. If necessary, learners could review or change their objective on the progress page of the course. A survey was conducted to gather further feedback about the perceived usefulness, next to the acceptance. All learners assigned to the groups 2, 3, or 4 could participate. A total number of 163 submissions was collected.

- In total, a considerable portion of learners (38.0%–47.7%) sets a personal learning objective for a course which demonstrates the interest to select a personal objective.
- In the groups where the selection modal was shown, 49.4%–58.7% of the learners selected an objective while the more subtle alternative of only showing an infobox attracted noticeably fewer learners (19.2%–32.8%).
- Hence, nudging learners with an objective modal while offering multiple possibilities to set an objective is identified as the best-suited approach to engage learners.

- The majority of learners intended to complete the course (65.7%–55.9%) while the minority of the learners either wanted to get an overview of the course or focused on a more specific aspect of the course (34.3%–44.1%).
- The interests and intentions for a course vary remarkably and learners do not solely focus on course completion but also prefer individual learning paths. Consequently, the concept of providing learning objectives based on dedicated topic units is reasonable and accepted by the learners and helps to identify a variety of intentions.
- These types of learning objectives can be applied best to courses with a wide subject breadth or depth to offer focus topics or proficiency levels to learners.
- Changes between objectives after an initial selection rarely happened (0.8%–2.7%).
- The provided features are well-perceived by the learners: 64.42% like to have learning objectives available in other courses as well and 54.61% consider the selection of an objective as useful.

In the next iteration, the learner dashboard was combined with personalized learning objectives. Thus, for the first time, learners were able to evaluate the progression and achievement of their learning objectives. In addition, the UI variant from the last experiment was implemented, which combines the modal with the infobox. In a last evaluation, platform data from two courses with a total of 2,642 'shows at middle' were analyzed to investigate which learners selected an objective regarding their socio-demographic and geographical background, and how successfully learners achieved their selected objectives. Also, a survey with 279 submissions was conducted in one course to compare the general course satisfaction of students with and without a selected learning objective. Here, we compared both user groups regarding statistically significant differences utilizing the Mann-Whitney U test and assessed the practical relevance of the descriptive statistics. Answers were collected from 163 students without a selected learning objective and 116 students with a selected learning objective.

- No practically relevant differences are found between students with selected learning objectives and the total course population regarding their age, gender, degree, career status, or location.
- Hence, students with a selected learning objective are no more, but also no less satisfied with the course than students without a selected learning objective. The generally very positive results, which are distorted by a proven survivorship bias, show no statistically significant differences between the two cohorts and no practical relevance is derived as well.
- Objective selection rates differ largely between courses (28.57%–63.87%), probably due to the fact that objectives are defined individually for each course by different teaching teams. This is supported by the findings of the last study.
- Learners enroll for courses with varying outcome intentions. Even though the most frequently selected learning objectives included a graded certificate (71.88%–83.67%), which is reasonable considering that the courses were primarily designed with this intention, also a notable amount of learners were not interested in gaining a Record of Achievement (16.33%–28.12%). This is also supported by the findings of the last study.

- Compared to typical MOOC completion rates, the objective achievement rates ranging from 18.18% up to 46.97% are considered a success. Also, between 6.06% and 18.18% of the learners even exceeded their objectives, which indicates an increase in motivation during the course.
- As this approach to measure success in MOOCs is new, only the certification rates for objectives that include a Record of Achievement can be compared with the traditional approach, which assumes that all learners aim to achieve a certificate.
- Improved certification rates of 11.62% and 12.63% are found, comparing gained Records of Achievement for the total course population and users who selected and completed objectives that included this certificate. We consider both rates as a practically significant improvement.

Research Question 3: How can data-driven insights support teaching in MOOCs?

To pursue this question, a learning analytics dashboard for teachers was developed and examined in two iterations. In particular, it is intended to support teaching teams in monitoring their courses with thousands of learners and responding to problems more effectively. The first version improved an existing prototype, which was technically and functionally refined based on identified shortcomings and requirements. Especially the performance and the hierarchy and structure of the visualizations and widgets were revised and web analytics incorporated. By means of a survey, the usability and usefulness were investigated and compared with the previous prototype, especially with regard to the use of WA methods. In total, 11 teaching experts participated in the survey. To evaluate statistically significant differences, a Wilcoxon test was performed and effect sizes were computed with Cohen's d.

- The revision of the dashboards has statistically significant differences regarding the ease of use and satisfaction with intermediate to large effect sizes. Especially the clearness is improved as shown by the quantitative, but also qualitative evaluation.
- No statistically significant differences are found regarding the usefulness of WA metrics.
- A large amount of the existing LA metrics, which are indeed proven to be useful for the stakeholders, can be realized with WA methods as well. Consequently, WA can still provide useful insights in the context of LA.
- WA can contribute to an increase of the dashboard's general performance by making use of the cloud infrastructure of WA tools. This is demonstrated by a statistically significant difference with a large effect size in terms of perceived performance.

The second version of the teacher dashboard incorporated the remaining requirements, which targeted the functional domains of the HPI MOOC Platform. Afterward, a last evaluation was conducted, which examined how the dashboard and statistics are used and perceived by teaching teams. For the first part, platform data from 33 courses and 4,461 dashboard and statistics page visits were analyzed. For the second part, another survey was conducted with 23 teaching experts. The survey questioned the usefulness of the various visualizations and statistics during different teaching phases of a course and the teachers' perception with different experience levels based on the EFLA questionnaire.

- The dashboard and statistics pages are used very frequently, between 8 and 687 times with an average of 135 hits per course.
- The dashboard and statistics pages are used by several teachers and managers, with 3 to 15 people involved and an average of 7 teachers and managers per course.
- The dashboard and statistics pages are used mainly while a course is running (51.5%) and afterward for evaluation (37%), but rather less during the design and development of a course (11.5%) probably due to the lack of available data.
- The high usage figures clearly show that learning analytics dashboards are a key tool and an integral part for teaching in MOOCs.
- The visualizations and widgets regarding enrollments and learners, certificates, geographical data (user locations), demographical information (age distribution), as well as quiz submissions, item visits, and video interactions, are considered most useful.
- The usefulness is perceived as rather limited in the pre-design and design and development phase (2.20–2.61 on a scale from 1 to 7), whereas a higher usefulness is perceived during the facilitation and evaluation phase (4.51–5.06). The results are in line with the usage statistics.
- The level of experience in teaching has only a minor impact on the perception, although teachers with little (EFLA Score: 69) or very much practice (65) in MOOCs find these tools slightly more helpful for their work than those with moderate experience (59).
- The greatest benefit of the dashboard and statistics' data-driven insights is the increased awareness of the students' learning situation (EFLA-AR1: 7.48).

5.2 Contributions of the Thesis

In addition to the presented findings which advance the body of knowledge in the relevant research disciplines, this thesis also elaborated a conceptual blueprint of how learning analytics can be technically integrated and applied in MOOCs to provide learners and teachers with data-driven insights. In summary, this resulted in four cornerstones that serve as a guideline and reference for related studies and projects.

- The Learning Analytics Architecture: A technical infrastructure to collect, process, and analyze event-driven learning data based on schema-agnostic pipelining in a service-oriented MOOC platform.
- The Learning Analytics Dashboard for Learners: A tool for data-driven support of self-regulated learning, in particular to enable learners to evaluate and plan their learning activities, progress, and success by themselves.
- Personalized Learning Objectives: A set of features to better connect learners' success to their personal intentions based on selected learning objectives to offer guidance and align the provided data-driven insights about their learning progress.

- The Learning Analytics Dashboard for Teachers: A tool supporting teachers with data-driven insights to enable the monitoring of their courses with thousands of learners, identify potential issues, and take informed action.

In parallel, the various evaluations of the presented case studies demonstrated how learning analytics can be used from a researcher's perspective. Therefore, the implemented LA capabilities enabled many innovative approaches to the research design, applying proven statistical methods to previously unavailable learning data. In particular, these help to better understand the learning and teaching in MOOCs. Examples include:

- An inferential statistical analysis of seven learning behavior metrics of learners who used the mobile applications next to the web platform and those who did not (Subsection 2.3.4).
- An inferential statistical analysis of two learning completion metrics of learners with regard to the use of three dashboard variants in an A/B/n test (Subsection 3.4.2).
- A descriptive statistical analysis of the achievement rates of self-reported learning objectives (Subsection 3.5.1).
- A descriptive statistical analysis of objective selection rates of learners with regard to the use of the three objective selection UI variants in an A/B/n test (Subsection 3.6.4).
- A descriptive statistical analysis of the achievement rates of personalized learning objectives (Subsection 3.7.2).
- A descriptive statistical analysis of the usage of the teacher dashboard and statistics pages in different teaching phases of the corresponding courses (Subsection 4.3.2).

5.3 Limitations of the Thesis

The elaborated findings and research contributions of this thesis are subject to certain technical and methodical limitations that have to be taken into account. First, the proposed LA architecture was designed specifically for the HPI MOOC Platform. At the time of completion of this thesis, the platform was still closed source, but the disclosure of the code base as open source has been already planned. Nevertheless, even then the specific implementations are not transferable to other MOOC platforms without major technical efforts. Also, the context of a service-oriented architecture involved many specific requirements. The most likely to be reusable is the dedicated learning analytics service, also as an extension for monolithic applications. All user-facing features were integrated directly into the existing frontend of the platform and are therefore strongly coupled to it. Hence, the technical prototypes are only transferable to other platforms on a conceptual level. The widgets and visualizations displayed on the dashboards were selected through expert interviews and related work, and refined with user feedback. Nevertheless, not all possible elements were implemented and tested. Thus, the presented results refer only to the selection we made. Also, the coarse selection of learning objectives defined and offered by the teaching teams can reveal more precise nuances regarding learners' intentions through finer granularity, which enables to offer more useful data-driven feedback.

Second, the conducted case studies were field experiments that were subject to uncontrollable effects in contrast to lab settings. This caused, for example, unequal sample sizes, which in conjunction with an overall high number of participants can lead quickly to statistically significant differences and affect the probability of Type II errors. Additionally, the demographics of the users on the platforms studied are very homogeneous. As demonstrated, these are often well-educated male lifelong learners with a technical affinity. Although this makes the generalization of the findings difficult, the value of conducting experiments in an authentic learning environment with real users outweighs these drawbacks. Furthermore, the collected self-reported data relied on the participants' self-awareness and truthfulness—and was always requested voluntarily. In addition, post-course surveys were subject to a demonstrated survivorship bias which shifted the results into a more positive direction. Besides, it was difficult to engage a large number of teachers as survey participants. Additional qualitative research methods can complement this. Lastly, all studies had a technical focus and were conducted in the context of MOOCs, which limits the applicability of the results to other learning settings such as blended or distance higher education and omits a deeper educational sciences perspective.

5.4 Implications and Future Research

The contributions of this thesis are most beneficial for the further development and research of innovative data-driven educational tools, as well as learning and teaching practices in MOOCs. We showed that self-regulated learning and in particular self-evaluation of students utilizing a learning analytics dashboard can be supported more successfully by incorporating the various intentions of lifelong learners as a reference. This was enabled by setting objectives, which in turn must be provided meaningfully by teaching teams and considered in the instructional design of courses. Hence, this is crucial for a large-scale adaption of our corresponding findings. Additionally, we demonstrated that a learning analytics dashboard is an essential tool for teachers to conduct and monitor their MOOCs. All in all, we provided further evidence that learning analytics methods are clearly suited to address the challenges caused by the massiveness of MOOCs. Our presented concepts are tested and applicable in practice.

However, learning analytics provides further approaches. As the research field evolves rapidly, additional user feedback can be implemented, and limitations be addressed, there are several opportunities to continue this work. To verify the external validity of the presented results, the research prototypes need to be transferred to other MOOC platforms and further VLEs. This enables the examination of additional learning settings and more diverse demographics. As already discussed, this involves major technical efforts. By fully supporting the xAPI standard, at least the integration of the LA service into other systems becomes easier when the code is published as open source.

The learner dashboard can be further enhanced and explored in terms of customization options, support for action, and additional reference frames. Learners expressed interest in turning widgets on and off, as well as reordering them. Customization is also discussed in current literature as a design recommendation for LADs [58], as well as more support for action [56]. We already provide information for planning, such as the estimated time effort for learning items and objectives, as well as objective-based guidance and recommendations for repetitions. Besides, recommendations for adapting the learning behavior can be explored to achieve a learning objective more efficiently and effectively. Furthermore, in addition to the already supported reference frames regarding progress and achievement, social reference frames can be examined. Nevertheless, social comparison with peers, such as low or high performers or learners with the same learning objective, have to be investigated with caution as this can also lead to negative effects [93, 109]. In general, it has to be ensured that all dashboard elements are based on learning theory and all phases of SRL are supported. Nevertheless, they have to be technically feasible at the same time. In this regard, current literature is concerned with approaches to better automate the detection of SRL activities [35, 128, 130] so that evaluations are less dependent on self-reported data which is also a limitation of this work.

For the concept of personalized learning objectives, it is of interest to define them across multiple courses, as well as to add a competency model. The implemented data model already supports this abstraction. Another approach is to enable learners to formulate learning objectives individually for more customization. However, this conflicts with the ability to automatically measure progress and achievement of learning objectives, which was one of our main requirements. Furthermore, the current concept is mainly focused on active courses during their runtime. However, learning objectives can also be beneficial for self-paced courses, which can be explored further.

There were also several suggestions for the continuation of the teacher dashboard. First, more interaction options need to be provided for the visualizations, as well as recommendations for action. For the latter, the initiated concept of automated quality control can be enhanced and further integrated [106]. Moreover, the usefulness of the teacher dashboard can be increased, especially in early course phases, by providing reference frames such as the comparison with other courses or the overall platform. Also, the design can be further informed by conceptual models such as teacher inquiry or data-driven instruction. However, this has to be practically feasible and depends on the instruction design and production process of a MOOC.

The outlined approaches to continue our work show that the potential of data-driven learning and teaching tools has not been fully explored yet. We are proud to have provided an essential contribution to this young research field and to have implemented practical tools that enhance learning and teaching in an increasingly digital world. We are therefore convinced that our research and findings support personal and societal development—as learning means development.

Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
cMOOC	Connectivism MOOC
CoP	Confirmation of Participation
DSL	Domain-Specific Language
EDM	Educational Data Mining
EFLA	Evaluation Framework for Learning Analytics
ETL	Extract, Transform, Load
EU	European Union
GDPR	General Data Protection Regulation
HPI	Hasso Plattner Institute
HTTP	Hypertext Transfer Protocol
KPI	Key Performance Indicator
LA	Learning Analytics
LAD	Learning Analytics Dashboard
LMS	Learning Management System
LRS	Learning Record Store
LTI	Learning Tools Interoperability
MIT	Massachusetts Institute of Technology
MLA	Mobile Learning Analytics
MOOC	Massive Open Online Course
OCW	Open Course Ware
OER	Open Educational Resources
OLAP	Online Analytical Processing
OLTP	Online Transaction Processing
PII	Personally Identifiable Information
PLO	Personalized Learning Objective
QC	Qualified Certificate

RDF	Resource Description Framework
RoA	Record of Achievement
SDK	Software Development Kit
SOA	Service-Oriented Architecture
SRL	Self-Regulated Learning
TAM	Technology Acceptance Model
TEL	Technology-Enhanced Learning
UI	User Interface
ULA	Ubiquitous Learning Analytics
URL	Uniform Resource Locator
VLE	Virtual Learning Environment
WA	Web Analytics
WHO	World Health Organization
xMOOC	Extension MOOC

References

1. G. D. Abowd, A. K. Dey, P. J. Brown, N. Davies, M. Smith, and P. Steggles. "Towards a Better Understanding of Context and Context-Awareness". In: *Handheld and Ubiquitous Computing*. Ed. by H.-W. Gellersen. Springer Berlin Heidelberg, 1999, pp. 304–307. DOI: 10.1007/3-540-48157-5_29.
2. N. R. Aljohani and H. C. Davis. "Learning Analytics in Mobile and Ubiquitous Learning Environments". In: *Proceedings of the 11th International Conference on Mobile and Contextual Learning*. mLearn '12. CEUR-WS.org, 2012, pp. 70–77. URL: http://ceur-ws.org/Vol-955/papers/paper_70.pdf.
3. M. E. Alonso-Mencía, C. Alario-Hoyos, J. Maldonado-Mahauad, I. Estévez-Ayres, M. Pérez-Sanagustín, and C. D. Kloos. "Self-regulated learning in MOOCs: lessons learned from a literature review". *Educational Review* 72:3, 2019, pp. 319–345. DOI: 10.1080/00131911.2019.1566208.
4. M. A. Andor, K. M. Fels, J. Renz, and S. Rzepka. "Do Planning Prompts Increase Educational Success? Evidence from Randomized Controlled Trials in MOOCs". *Ruhr Economic Papers* 790, 2018. DOI: 10.4419/86788918.
5. M. Arrigo, G. Fulantelli, and D. Taibi. "Challenges of Using Learning Analytics Techniques to Support Mobile Learning". *International Association for Development of the Information Society*, 2015. URL: https://eric.ed.gov/?id=ED562435.
6. R. Baker. "Stupid Tutoring Systems, Intelligent Humans". *International Journal of Artificial Intelligence in Education* 26, 2016, pp. 600–614. DOI: 10.1007/s40593-016-0105-0.
7. K. Bhageshpur. *Data Is The New Oil – And That's A Good Thing*. 2019. URL: https://www.forbes.com/sites/forbestechcouncil/2019/11/15/data-is-the-new-oil-and-thats-a-good-thing/ (visited on 03/22/2021).
8. R. Bodily and K. Verbert. "Trends and Issues in Student-Facing Learning Analytics Reporting Systems Research". In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. Association for Computing Machinery, pp. 309–318. DOI: 10.1145/3027385.3027403.
9. M. Bothe and C. Meinel. "Applied Mobile-Assisted Seamless Learning Techniques in MOOCs". In: *Digital Education: At the MOOC Crossroads Where the Interests of Academia and Business Converge*. EMOOCs '19. Springer International Publishing, 2019, pp. 21–30. DOI: 10.1007/978-3-030-19875-6_3.
10. M. Bothe, J. Renz, T. Rohloff, and C. Meinel. "From MOOCs to Micro Learning Activities". In: *Proceedings of the 2019 IEEE Global Engineering Education Conference*. EDUCON '19. 2019, pp. 280–288. DOI: 10.1109/EDUCON.2019.8725043.
11. L. B. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho, and D. T. Seaton. "Studying learning in the worldwide classroom: Research into edX's first MOOC". *Research & Practice in Assessment* 8, 2013, pp. 13–25. URL: http://www.rpajournal.com/dev/wp-content/uploads/2013/05/SF2.pdf.
12. D. Bri, M. García, H. Coll, and J. Lloret. "A Study of Virtual Learning Environments". *WSEAS Transactions on Advances in Engineering Education* 6:1, 2009, pp. 33–43.

13. J. Broadbent and W. Poon. "Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review". *The Internet and Higher Education*, 2015. DOI: 10.1016/j.iheduc.2015.04.007.

14. E. Brynjolfsson, L. M. Hitt, and H. H. Kim. "Strength in Numbers: How Does Data-Driven Decisionmaking Affect Firm Performance?" *SSRN eLibrary*, 2011. DOI: 10.2139/ssrn.1819486.

15. N. E. Cagiltay, K. Cagiltay, and B. Celik. "An Analysis of Course Characteristics, Learner Characteristics, and Certification Rates in MITx MOOCs". *International Review of Research in Open and Distributed Learning* 21:3, 2020, pp. 121–139. DOI: 10.19173/irrodl.v21i3.4698.

16. S. Charleer, J. Klerkx, E. Duval, T. De Laet, and K. Verbert. "Creating Effective Learning Analytics Dashboards: Lessons Learnt". In: *Adaptive and Adaptable Learning*. EC-TEL '16. Springer International Publishing, 2016, pp. 42–56. DOI: 10.1007/978-3-319-45153-4_4.

17. M. A. Chatti, V. Lukarov, H. Thüs, A. Muslim, F. A. M. Yousef, U. Wahid, C. Greven, A. Chakrabarti, and U. Schroeder. "Learning Analytics: Challenges and Future Research Directions". *e-learning and education (eleed)* 10:1, 2014. URL: http://nbn-resolving.de/urn:nbn:de:0009-5-40350.

18. X. Che, H. Yang, and C. Meinel. "Adaptive E-Lecture Video Outline Extraction Based on Slides Analysis". In: *Advances in Web-Based Learning*. CWL '15. Springer International Publishing, 2015, pp. 59–68. DOI: 10.1007/978-3-319-25515-6_6.

19. A. Cooper. *Learning Analytics Interoperability - The Big Picture In Brief*. Technical report. 2014. URL: http://laceproject.eu/publications/briefing-01.pdf.

20. K. Cotton. "Monitoring Student Learning in the Classroom". *School Improvement Research Series*, 4 1988. URL: https://eric.ed.gov/?id=ED298085.

21. F. Dalipi, A. S. Imran, and Z. Kastrati. "MOOC dropout prediction using machine learning techniques: Review and research challenges". In: *Proceedings of the 2018 IEEE Global Engineering Education Conference*. EDUCON '18. IEEE, 2018, pp. 1007–1014. DOI: 10.1109/EDUCON.2018.8363340.

22. J. Daniel. "Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility". *Journal of Interactive Media in Education* 2012:3, 2012. DOI: 10.5334/2012-18.

23. D. Davis, G. Chen, I. Jivet, C. Hauff, and G. Houben. "Encouraging Metacognition & Self-Regulation in MOOCs through Increased Learner Feedback". In: *Proceedings of the LAK 2016 Workshop on Learning Analytics for Learners*. CEUR-WS.org, 2016, pp. 17–22. URL: http://ceur-ws.org/Vol-1596/paper3.pdf.

24. D. Davis, G. Chen, T. van der Zee, C. Hauff, and G.-J. Houben. "Retrieval Practice and Study Planning in MOOCs: Exploring Classroom-Based Self-regulated Learning Strategies at Scale". In: *Adaptive and Adaptable Learning*. EC-TEL '16. Springer International Publishing, 2016, pp. 57–71. DOI: 10.1007/978-3-319-45153-4_5.

25. D. Davis, I. Jivet, R. F. Kizilcec, G. Chen, C. Hauff, and G.-J. Houben. "Follow the Successful Crowd: Raising MOOC Completion Rates Through Social Comparison at Scale". In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. Association for Computing Machinery, 2017, pp. 454–463. DOI: 10.1145/3027385.3027411.

26. F. D. Davis. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology". *MIS Quarterly* 13:3, 1989, pp. 319–340. DOI: 10.2307/249008.

27. S. Dawson, S. Joksimovic, O. Poquet, and G. Siemens. "Increasing the Impact of Learning Analytics". In: *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*. LAK '19. Association for Computing Machinery, 2019, pp. 446–455. DOI: 10.1145/3303772.3303784.

28. P. Dillenbourg, D. Schneider, and P. Synteta. "Virtual Learning Environments". In: *Proceedings of the 3rd Hellenic Conference on Information & Communication Technologies in Education*. 2002, pp. 3–18.

29. G. T. Doran. "There's a SMART way to write management's goals and objectives". *Management Review* 70:11, 1981, pp. 35–36.

30. H. Drachsler and W. Greller. "Privacy and Analytics: It's a DELICATE Issue a Checklist for Trusted Learning Analytics". In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. LAK '16. Association for Computing Machinery, 2016, pp. 89–98. DOI: 10.1145/2883851.2883893.

31. H. Drachsler and W. Greller. "The Pulse of Learning Analytics Understandings and Expectations from the Stakeholders". In: *Proceedings of the 2nd International Conference on Learning Analytics & Knowledge*. LAK '12. Association for Computing Machinery, 2012, pp. 120–129. DOI: 10.1145/2330601.2330634.

32. H. Drachsler and M. Kalz. "The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges". *Journal of Computer Assisted Learning* 32:3, 2016, pp. 281–290. DOI: 10.1111/jcal.12135.

33. edX Inc. *Open edX Developer's Guide - Analytics*. URL: https://edx.readthedocs.io/projects/edx-developer-guide/en/latest/analytics.html (visited on 07/03/2018).

34. E. J. Emanuel. "Online education: MOOCs taken by educated few". *Nature* 503:7476, 2013, p. 342. URL: https://www.nature.com/articles/503342a.

35. Y. Fan, J. Saint, S. Singh, J. Jovanovic, and D. Gašević. "A Learning Analytic Approach to Unveiling Self-Regulatory Processes in Learning Tactics". In: *Proceedings of the 11th International Conference on Learning Analytics and Knowledge*. LAK '21. Association for Computing Machinery, 2021, pp. 184–195. DOI: 10.1145/3448139.3448211.

36. M. A. Farsani, M. Beikmohammadi, and A. Mohebbi. "Self-Regulated Learning, Goal-Oriented Learning, and Academic Writing Performance of Undergraduate Iranian EFL Learners". *TESL-EJ* 18:2, 2014. URL: https://eric.ed.gov/?id=EJ1045129.

37. R. Ferguson. "Learning analytics: drivers, developments and challenges". *International Journal of Technology Enhanced Learning* 4:5/6, 2012, pp. 304–317. DOI: 10.1504/IJTEL.2012.051816.

38. R. Ferguson, D. Clow, L. Macfadyen, A. Essa, S. Dawson, and S. Alexander. "Setting Learning Analytics in Context: Overcoming the Barriers to Large-Scale Adoption". In: *Proceedings of the Fourth International Conference on Learning Analytics & Knowledge*. LAK '14. Association for Computing Machinery, 2014, pp. 251–253. DOI: 10.1145/2567574.2567592.

39. S. Few. *Information Dashboard Design: The Effective Visual Communication of Data*. O'Reilly Media, Inc., 2006. ISBN: 0596100167.

40. A. Fox, M. A. Hearst, and M. T. Chi. "Chairs' Welcome Message". In: *Proceedings of the First ACM Conference on Learning @ Scale*. L@S '14. Association for Computing Machinery, 2014, pp. iii–iv. DOI: 10.1145/2556325.

41. C. Fredericks, G. Lopez, V. Shnayder, S. Rayyan, and D. Seaton. "Instructor Dashboards In EdX". In: *Proceedings of the Third ACM Conference on Learning @ Scale*. L@S '16. Association for Computing Machinery, 2016, pp. 335–336. DOI: 10.1145/2876034.2893405.

42. J. Galusha. "Barriers to Learning in Distance Education", 1998. URL: https://eric.ed.gov/?id=ED416377.

43. Gartner, Inc. *Gartner Information Technology Glossary: Dashboards*. 2020. URL: https://www.gartner.com/en/information-technology/glossary/dashboard (visited on 10/26/2020).
44. P. Goodyear. "Psychological Foundations for Networked Learning". In: *Networked Learning: Perspectives and Issues*. Springer, London, 2002, pp. 49–75. DOI: 10.1007/978-1-4471-0181-9_4.
45. Google Inc. *Measurement Protocol Reference*. URL: https://developers.google.com/analytics/devguides/collection/protocol/v1/reference (visited on 07/03/2018).
46. C. Grella and C. Meinel. "MOOCs as a Promoter of Gender Diversity in STEM?" In: *Proceedings of the 12th International Scientific Conference eLearning and Software for Education*. 2016. DOI: 10.12753/2066-026X-19-164.
47. W. Greller and H. Drachsler. "Translating Learning into Numbers: A Generic Framework for Learning Analytics". *Journal of Educational Technology & Society* 15:3, 2012, pp. 42–57. URL: https://www.jstor.org/stable/jeductechsoci.15.3.42.
48. T. Hecking, S. Manske, L. Bollen, S. Govaerts, A. Vozniuk, and H. U. Hoppe. "A Flexible and Extendable Learning Analytics Infrastructure". In: *Advances in Web-Based Learning*. ICWL '14. Springer International Publishing, 2014, pp. 123–132. DOI: 10.1007/978-3-319-09635-3_13.
49. M. Henderikx, K. Kreijns, and M. Kalz. "A Classification of Barriers that Influence Intention Achievement in MOOCs". In: *Lifelong Technology-Enhanced Learning*. EC-TEL '18. Springer International Publishing, 2018, pp. 3–15. DOI: 10.1007/978-3-319-98572-5_1.
50. M. A. Henderikx, K. Kreijns, and M. Kalz. "Refining success and dropout in massive open online courses based on the intention–behavior gap". *Distance Education* 38:3, 2017, pp. 353–368. DOI: 10.1080/01587919.2017.1369006.
51. C. Herodotou, B. Rienties, A. Boroowa, Z. Zdrahal, M. Hlosta, and G. Naydenova. "Implementing Predictive Learning Analytics on a Large Scale: The Teacher's Perspective". In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. Association for Computing Machinery, 2017, pp. 267–271. DOI: 10.1145/3027385.3027397.
52. P. Hill. *Emerging Student Patterns in MOOCs: A (Revised) Graphical View*. 2013. URL: https://eliterate.us/emerging-student-patterns-in-moocs-a-revised-graphical-view/ (visited on 01/28/2021).
53. H.-M. Huang. "Toward constructivism for adult learners in online learning environments". *British Journal of Educational Technology* 33:1, 2002, pp. 27–37. DOI: 10.1111/1467-8535.00236.
54. E. İnan and M. Ebner. "Learning Analytics and MOOCs". In: *Learning and Collaboration Technologies. Designing, Developing and Deploying Learning Experiences*. Springer International Publishing, 2020, pp. 241–254. DOI: 10.1007/978-3-030-50513-4_18.
55. P. Isaias and A. Backx Noronha Viana. "On the Design of a Teachers' Dashboard: Requirements and Insights". In: *Learning and Collaboration Technologies. Designing, Developing and Deploying Learning Experiences*. Springer International Publishing, 2020, pp. 255–269. DOI: 10.1007/978-3-030-50513-4_19.
56. I. Jivet, M. Scheffel, M. Schmitz, S. Robbers, M. Specht, and H. Drachsler. "From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education". *The Internet and Higher Education* 47, 2020, p. 100758. DOI: 10.1016/j.iheduc.2020.100758.
57. I. Jivet, M. Scheffel, M. Specht, and H. Drachsler. "License to Evaluate: Preparing Learning Analytics Dashboards for Educational Practice". In: *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*. LAK '18. Association for Computing Machinery, 2018, pp. 31–40. DOI: 10.1145/3170358.3170421.

58. I. Jivet, J. Wong, M. Scheffel, M. Valle Torre, M. Specht, and H. Drachsler. “Quantum of Choice: How Learners’ Feedback Monitoring Decisions, Goals and Self-Regulated Learning Skills Are Related”. In: *Proceedings of the 11th International Conference on Learning Analytics and Knowledge*. LAK ’21. Association for Computing Machinery, 2021, pp. 416–427. DOI: 10.1145/3448139.3448179.

59. C. John, T. Staubitz, and C. Meinel. “Performance of Men and Women in Graded Team Assignments in MOOCs”. In: *Proceedings of the 2019 IEEE Learning With MOOCS*. LWMOOCS ’19. 2019, pp. 30–35. DOI: 10.1109/LWMOOCS47620.2019.8939654.

60. C. John, T. Staubitz, and C. Meinel. “Took a MOOC. Got a Certificate. What now?” In: *Proceedings of the 2019 IEEE Frontiers in Education Conference*. FIE ’19. IEEE, 2019, pp. 1–9. DOI: 10.1109/FIE43999.2019.9028698.

61. K. Jordan. “Initial Trends in Enrolment and Completion of Massive Open Online Courses”. *The International Review of Research in Open and Distributed Learning* 15:1, 2014. DOI: 10.19173/irrodl.v15i1.1651.

62. K. Jordan. “Massive Open Online Course Completion Rates Revisited: Assessment, Length And Attrition”. *International Review of Research in Open and Distributed Learning* 16:3, 2015. DOI: 10.19173/irrodl.v16i3.2112.

63. D. Joyner. “Congruency, Adaptivity, Modularity, and Personalization: Four Experiments in Teaching Introduction to Computing”. In: *Proceedings of the Fourth ACM Conference on Learning @ Scale*. L@S ’17. Association for Computing Machinery, 2017, pp. 307–310. DOI: 10.1145/3051457.3054011.

64. M. Kalz. “Lifelong Learning and Its Support with New Technologies”. In: *International Encyclopedia of the Social & Behavioral Sciences (Second Edition)*. Ed. by J. D. Wright. Elsevier, 2015, pp. 93–99. DOI: 10.1016/B978-0-08-097086-8.92006-3.

65. R. F. Kizilcec, C. Piech, and E. Schneider. “Deconstructing Disengagement: Analyzing Learner Subpopulations in Massive Open Online Courses”. In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge*. LAK ’13. Association for Computing Machinery, 2013, pp. 170–179. DOI: 10.1145/2460296.2460330.

66. R. F. Kizilcec, A. Saltarelli, J. Reich, and G. Cohen. “Closing global achievement gaps in MOOCs”. *Science* 355:6322, 2017, pp. 251–252. DOI: 10.1126/science.aag2063.

67. R. F. Kizilcec and A. J. Saltarelli. “Can a Diversity Statement Increase Diversity in MOOCs?” In: *Proceedings of the Sixth ACM Conference on Learning @ Scale*. L@S ’19. Association for Computing Machinery, 2019, pp. 1–8. DOI: 10.1145/3330430.3333633.

68. R. F. Kizilcec and E. Schneider. “Motivation as a Lens to Understand Online Learners: Toward Data-Driven Design with the OLEI Scale”. *ACM Transactions on Computer-Human Interaction* 22, 2015, 6:1–6:24. DOI: 10.1145/2699735.

69. R. F. Kizilcec and S. Halawa. “Attrition and Achievement Gaps in Online Learning”. In: *Proceedings of the Second ACM Conference on Learning @ Scale*. L@S ’15. Association for Computing Machinery, 2015, pp. 57–66. DOI: 10.1145/2724660.2724680.

70. R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado. “Recommending Self-Regulated Learning Strategies Does Not Improve Performance in a MOOC”. In: *Proceedings of the Third ACM Conference on Learning @ Scale*. L@S ’16. Association for Computing Machinery, 2016, pp. 101–104. DOI: 10.1145/2876034.2893378.

71. R. F. Kizilcec, M. Pérez-Sanagustín, and J. J. Maldonado. "Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses". *Computers & Education* 104, 2017, pp. 18–33. DOI: 10.1016/j.compedu.2016.10.001.

72. J. Klerkx, K. Verbert, and E. Duval. "Learning Analytics Dashboards". In: *The Handbook of Learning Analytics*. Ed. by C. Lang, G. Siemens, A. F. Wise, and D. Gaševic. 1st ed. Society for Learning Analytics Research (SoLAR), 2017, pp. 143–150. DOI: 10.18608/hla17.012.

73. V. Kovanović, M. Scheffel, N. Pinkwart, and K. Verbert. "LAK 2020 Program Chairs' Welcome". In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. LAK '20. Association for Computing Machinery, 2020. DOI: 10.1145/3375462.

74. D. R. Krathwohl. "A Revision of Bloom's Taxonomy: An Overview". *Theory Into Practice* 41:4, 2002, pp. 212–218. DOI: 10.1207/s15430421tip4104_2.

75. S. P. Lajoie and R. Azevedo. "Teaching and Learning in Technology-Rich Environments". In: *Handbook of Educational Psychology*. Routledge, 2006. DOI: 10.4324/9780203874790.ch35.

76. D. Lee, S. Watson, and W. Watson. "Systematic literature review on self-regulated learning in massive open online courses". *Australasian Journal of Educational Technology*, 2018. DOI: 10.14742/ajet.3749.

77. M. Leon, R. Cobos, K. Dickens, S. White, and H. Davis. "Visualising the MOOC experience: a dynamic MOOC dashboard built through institutional collaboration". In: *Proceedings of the European Stakeholder Summit on experiences and best practices in and around MOOCs*. EMOOCs '16. BoD, 2016, pp. 461–469.

78. A. Littlejohn. "Understanding Massive Open Online Courses". *CEMCA EdTech Notes*, 2013.

79. A. Littlejohn and C. Milligan. "Designing MOOCs for professional learners: Tools and patterns to encourage self-regulated learning". *eLearning Papers* 42, 2015, pp. 38–45. URL: http://oro.open.ac.uk/46385/.

80. T. R. Liyanagunawardena, P. Parslow, and S. A. Williams. "Dropout: MOOC Participants' Perspective". In: *The Second MOOC European Stakeholders Summit*. EMOOCs 2014. 2014, pp. 95–100. URL: http://centaur.reading.ac.uk/36002/.

81. H. Luo, S. Rocco, and C. Schaad. "Using Google Analytics to Understand Online Learning: A Case Study of a Graduate-Level Online Course". In: *2015 International Conference of Educational Innovation through Technology*. EITT '15. IEEE, 2015, pp. 264–268. DOI: 10.1109/EITT.2015.62.

82. A. Margaryan, M. Bianco, and A. Littlejohn. "Instructional quality of Massive Open Online Courses (MOOCs)". 80, 2015, pp. 77–83. DOI: 10.1016/j.compedu.2014.08.005.

83. C. Marinagi, C. Skourlas, and P. Belsis. "Employing Ubiquitous Computing Devices and Technologies in the Higher Education Classroom of the Future". *Procedia - Social and Behavioral Sciences* 73, 2013, pp. 487–494. DOI: https://doi.org/10.1016/j.sbspro.2013.02.081.

84. R. Martinez-Maldonado, A. Pardo, N. Mirriahi, K. Yacef, J. Kay, and A. Clayphan. "LATUX: an Iterative Workflow for Designing, Validating and Deploying Learning Analytics Visualisations". *Journal of Learning Analytics* 2:3, 2016, pp. 9–39. DOI: 10.18608/jla.2015.23.3.

85. T. Mayes and S. De Freitas. "Review of E-Learning Theories, Frameworks and Models". *JISC E-Learning Models Desk Study*, 2004. URL: http://www.jisc.ac.uk/whatwedo/programmes/elearningpedagogy/outcomes.aspx.

86. R. Mazza and V. Dimitrova. "Visualising Student Tracking Data to Support Instructors in Web-Based Distance Education". In: *Proceedings of the 13th International World Wide Web Conference on Alternate Track Papers & Posters*. WWW Alt. '04. Association for Computing Machinery, 2004, pp. 154–161. DOI: 10.1145/1013367.1013393.

87. C. Meinel and C. Willems. *openHPI: The MOOC Offer at Hasso Plattner Institute*. Technical report. University of Potsdam, 2013. URL: `https://nbn-resolving.org/urn:nbn:de:kobv:517-opus-67176`.

88. C. Milligan and A. Littlejohn. "How health professionals regulate their learning in massive open online courses". *The Internet and Higher Education* 31, 2016, pp. 113–121. DOI: `10.1016/j.iheduc.2016.07.005`.

89. C. Milligan and A. Littlejohn. "Why Study on a MOOC? The Motives of Students and Professionals". *The International Review of Research in Open and Distributed Learning* 18:2, 2017. DOI: `10.19173/irrodl.v18i2.3033`.

90. MIT Office of Educational Innovation and Technology and New Media Consortium. *Online Course Design Guide*. URL: `https://dltoolkit.mit.edu/online-course-design-guide/` (visited on 01/22/2021).

91. B. Moissa, L. S. de Carvalho, and I. Gasparini. "A Web Analytics and Visualization Tool to Understand Students' Behavior in an Adaptive E-Learning System". In: *Learning and Collaboration Technologies. Designing and Developing Novel Learning Experiences*. LCT '14. Springer International Publishing, 2014, pp. 312–321. DOI: `10.1007/978-3-319-07482-5_30`.

92. A. Mottus, Kinshuk, S. Graf, and N.-S. Chen. "Use of Dashboards and Visualization Techniques to Support Teacher Decision Making". In: *Ubiquitous Learning Environments and Technologies*. Ed. by Kinshuk and R. Huang. Springer Berlin Heidelberg, 2015, pp. 181–199. DOI: `10.1007/978-3-662-44659-1_10`.

93. D. Muller and M.-P. Fayant. "On Being Exposed to Superior Others: Consequences of Self-Threatening Upward Social Comparisons". *Social and Personality Psychology Compass* 4:8, 2010, pp. 621–634. DOI: `10.1111/j.1751-9004.2010.00279.x`.

94. K. S. Na and Z. Tasir. "Identifying at-risk students in online learning by analysing learning behaviour: A systematic review". In: *Proceedings of the 2017 IEEE Conference on Big Data and Analytics*. ICBDA '17. 2017, pp. 118–123. DOI: `10.1109/ICBDAA.2017.8284117`.

95. D. Onah and J. Sinclair. "Assessing Self-Regulation of Learning Dimensions in a Stand-alone MOOC Platform". *International Journal of Engineering Pedagogy* 7:2, 2017, pp. 4–21. DOI: `10.3991/ijep.v7i2.6511`.

96. L. Pappano. *The Year of the MOOC*. 2012. URL: `http://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html` (visited on 11/12/2020).

97. T. Park, H. Cha, and H. Lee. "A Study on Design Guidelines of Learning Analytics to Facilitate Self-Regulated Learning in MOOCs". *Educational Technology International* 17:1, 2016, pp. 117–150. URL: `http://www.kset.or.kr/eti_ojs/index.php/instruction/article/view/61`.

98. R. A. Pérez-Álvarez, J. Maldonado-Mahauad, K. Sharma, D. Sapunar-Opazo, and M. Pérez-Sanagustín. "Characterizing Learners' Engagement in MOOCs: An Observational Case Study Using the NoteMyProgress Tool for Supporting Self-Regulation". *IEEE Transactions on Learning Technologies* 13, 2020, pp. 676–688. DOI: `10.1109/TLT.2020.3003220`.

99. P. R. Pintrich. "The Role of Goal Orientation in Self-Regulated Learning". In: *Handbook of Self-Regulation*. Ed. by M. Boekaerts, P. R. Pintrich, and M. Zeidner. Academic Press, 2000, pp. 451–502. DOI: `10.1016/B978-012109890-2/50043-3`.

100. D. Pérez-Berenguer, M. Kessler, and J. García-Molina. "A Customizable and Incremental Processing Approach for Learning Analytics". *IEEE Access* 8, 2020, pp. 36350–36362. DOI: `10.1109/ACCESS.2020.2975384`.

101. T. Rabelo, M. Lama, R. R. Amorim, and J. C. Vidal. "SmartLAK: A big data architecture for supporting learning analytics services". In: *Proceedings of the 2015 IEEE Frontiers in Education Conference*. FIE '15. IEEE, 2015, pp. 1–5. DOI: 10.1109/FIE.2015.7344147.

102. J. Reich. *MOOC Completion and Retention in the Context of Student Intent*. 2014. URL: https://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent (visited on 01/28/2021).

103. J. Reich and J. A. Ruipérez-Valiente. "The MOOC pivot". *Science* 363:6423, 2019, pp. 130–131. DOI: 10.1126/science.aav7958.

104. J. Renz, D. Hoffmann, T. Staubitz, and C. Meinel. "Using A/B Testing in MOOC Environments". In: *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge*. LAK '16. Association for Computing Machinery, 2016, pp. 304–313. DOI: 10.1145/2883851.2883876.

105. J. Renz, G. Navarro-Suarez, R. Sathi, T. Staubitz, and C. Meinel. "Enabling Schema Agnostic Learning Analytics in a Service-Oriented MOOC Platform". In: *Proceedings of the Third ACM Conference on Learning @ Scale*. L@S '16. Association for Computing Machinery, 2016, pp. 137–140. DOI: 10.1145/2876034.2893389.

106. J. Renz, T. Rohloff, and C. Meinel. "Automated Quality Assurance in MOOCs through Learning Analytics". In: *Joint Proceedings of the Pre-Conference Workshops of DeLFI and GMW 2017*. CEUR-WS.org, 2017. URL: http://ceur-ws.org/Vol-2092/paper24.pdf.

107. J. Renz, F. Schwerer, and C. Meinel. "openSAP: Evaluating xMOOC Usage and Challenges for Scalable and Open Enterprise Education". *International Journal of Advanced Corporate Learning* 9:2, 2016, e19076. DOI: 10.3991/ijac.v9i2.6008.

108. B. Rienties, C. Herodotou, T. Olney, M. Schencks, and A. Boroowa. "Making Sense of Learning Analytics Dashboards: A Technology Acceptance Perspective of 95 Teachers". *International Review of Research in Open and Distributed Learning* 19:5, 2018. DOI: 10.19173/irrodl.v19i5.3493.

109. T. Rogers and A. Feller. "Discouraged by Peer Excellence: Exposure to Exemplary Peer Performance Causes Quitting". *Psychological Science* 27:3, 2016, pp. 365–374. DOI: 10.1177/0956797615623770.

110. T. Rohloff, M. Bothe, J. Renz, and C. Meinel. "Towards a Better Understanding of Mobile Learning in MOOCs". In: *Proceedings of 2018 IEEE Learning With MOOCS*. LWMOOCS '18. IEEE, 2018, pp. 1–4. DOI: 10.1109/LWMOOCS.2018.8534685.

111. T. Rohloff and C. Meinel. "Towards Personalized Learning Objectives in MOOCs". In: *Lifelong Technology-Enhanced Learning*. EC-TEL '18. Springer International Publishing, 2018, pp. 202–215. DOI: 10.1007/978-3-319-98572-5_16.

112. T. Rohloff, S. Oldag, J. Renz, and C. Meinel. "Utilizing Web Analytics in the Context of Learning Analytics for Large-Scale Online Learning". In: *Proceedings of the 2019 IEEE Global Engineering Education Conference*. EDUCON '19. IEEE, 2019, pp. 296–305. DOI: 10.1109/EDUCON.2019.8725118.

113. T. Rohloff, J. Renz, M. Bothe, and C. Meinel. "Supporting Multi-Device E-Learning Patterns with Second Screen Mobile Applications". In: *Proceedings of the 16th World Conference on Mobile and Contextual Learning*. mLearn '17. Association for Computing Machinery, 2017, 25:1–25:8. DOI: 10.1145/3136907.3136931.

114. T. Rohloff, J. Renz, G. N. Suarez, and C. Meinel. "A Ubiquitous Learning Analytics Architecture for a Service-Oriented MOOC Platform". In: *Digital Education: At the MOOC Crossroads Where the Interests of Academia and Business Converge*. EMOOCs '19. Springer International Publishing, 2019, pp. 162–171. DOI: 10.1007/978-3-030-19875-6_19.

115. T. Rohloff, D. Sauer, and C. Meinel. “On the Acceptance and Usefulness of Personalized Learning Objectives in MOOCs”. In: *Proceedings of the Sixth ACM Conference on Learning @ Scale*. L@S ’19. Association for Computing Machinery, 2019, 4:1–4:10. DOI: 10.1145/3330430.3333624.

116. T. Rohloff, D. Sauer, and C. Meinel. “Student Perception of a Learner Dashboard in MOOCs to Encourage Self-Regulated Learning”. In: *Proceedings of the 2019 IEEE International Conference on Engineering, Technology and Education*. TALE ’19. IEEE, 2019, pp. 1–8. DOI: 10.1109/TALE48000.2019.9225939.

117. T. Rohloff, D. Sauer, and C. Meinel. “Students’ Achievement of Personalized Learning Objectives in MOOCs”. In: *Proceedings of the Seventh ACM Conference on Learning @ Scale*. L@S ’20. Association for Computing Machinery, 2020, pp. 147–156. DOI: 10.1145/3386527.3405918.

118. T. Rohloff, K. von Schmieden, and C. Meinel. “Students’ Satisfaction of a Design Thinking MOOC with Personalized Learning Objectives”. In: *Proceedings of 2020 IEEE Learning With MOOCs*. LWMOOCS ’20. IEEE, 2020, pp. 37–41. DOI: 10.1109/LWMOOCS50143.2020.9234349.

119. T. Rohloff, F. Schwerer, N. Schenk, and C. Meinel. “openSAP: Learner Behavior and Activity in Self-Paced Enterprise MOOCs”. *International Journal of Advanced Corporate Learning* 13:2, 2020, pp. 30–40. DOI: 10.3991/ijac.v13i2.16531.

120. T. Rohloff, H. Utunen, J. Renz, Y. Zhao, G. Gamhewage, and C. Meinel. “OpenWHO: Integrating Online Knowledge Transfer into Health Emergency Response”. In: *EC-TEL Practitioner Proceedings 2018*. EC-TEL ’18. CEUR-WS.org, 2018. URL: http://ceur-ws.org/Vol-2193/paper3.pdf.

121. B. Romanowski and A. Konak. “Using Google Analytics to Improve the Course Website of a Database Course”. In: *Proceedings of the 2016 ASEE Mid-Atlantic Section Conference*. American Society for Engineering Education, 2016. URL: https://www.hofstra.edu/pdf/academics/colleges/seas/asee-fall-2016/asee-midatlantic-f2016-konak.pdf.

122. C. Romero and S. Ventura. “Educational data mining and learning analytics: An updated survey”. *WIREs Data Mining and Knowledge Discovery* 10:3, 2020, e1355. DOI: https://doi.org/10.1002/widm.1355.

123. J. A. Ruipérez-Valiente, M. Jenner, T. Staubitz, X. Li, T. Rohloff, S. Halawa, C. Turro, Y. Cheng, J. Zhang, I. Despujol, and J. Reich. “Macro MOOC Learning Analytics: Exploring Trends across Global and Regional Providers”. In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. LAK ’20. Association for Computing Machinery, 2020, pp. 518–523. DOI: 10.1145/3375462.3375482.

124. J. A. Ruipérez-Valiente, P. J. Muñoz Merino, and C. D. Kloos. “An Architecture for Extending the Learning Analytics Support in the Khan Academy Framework”. In: *Proceedings of the First International Conference on Technological Ecosystem for Enhancing Multiculturality*. TEEM ’13. Association for Computing Machinery, 2013, pp. 277–284. DOI: 10.1145/2536536.2536578.

125. J. A. Ruipérez-Valiente, P. J. Muñoz-Merino, D. Leony, and C. Delgado Kloos. “ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform”. *Computers in Human Behavior* 47, 2015, pp. 139–148. DOI: 10.1016/j.chb.2014.07.002.

126. J. S. Ruiz, H. J. P. Díaz, J. A. Ruipérez-Valiente, P. J. Muñoz Merino, and C. D. Kloos. “Towards the Development of a Learning Analytics Extension in Open EdX”. In: *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality*. TEEM ’14. Association for Computing Machinery, 2014, pp. 299–306. DOI: 10.1145/2669711.2669914.

127. Rustici Software LLC. *What is the Experience API?* 2020. URL: https://xapi.com/overview/ (visited on 11/25/2020).

128. J. Saint, Y. Fan, S. Singh, D. Gasevic, and A. Pardo. "Using Process Mining to Analyse Self-Regulated Learning: A Systematic Analysis of Four Algorithms". In: *Proceedings of the 11th International Conference on Learning Analytics and Knowledge*. LAK '21. Association for Computing Machinery, 2021, pp. 333–343. DOI: 10.1145/3448139.3448171.

129. K. Saks and Äli Leijen. "Distinguishing Self-directed and Self-regulated Learning and Measuring them in the E-learning Context". *Procedia - Social and Behavioral Sciences* 112, 2014, pp. 190–198. DOI: 10.1016/j.sbspro.2014.01.1155.

130. F. Salehian Kia, M. Hatala, R. S. Baker, and S. D. Teasley. "Measuring Students' Self-Regulatory Phases in LMS with Behavior and Real-Time Self Report". In: *Proceedings of the 11th International Conference on Learning Analytics and Knowledge*. LAK '21. Association for Computing Machinery, 2021, pp. 259–268. DOI: 10.1145/3448139.3448164.

131. M. Scheffel. "The Evaluation Framework for Learning Analytics". PhD thesis. The Open University of the Netherlands, 2017. ISBN: 978-94-92231-81-9.

132. M. Scheffel, H. Drachsler, C. Toisoul, S. Ternier, and M. Specht. "The Proof of the Pudding: Examining Validity and Reliability of the Evaluation Framework for Learning Analytics". In: *Data Driven Approaches in Digital Education*. EC-TEL '17. Springer International Publishing, 2017, pp. 194–208. DOI: 10.1007/978-3-319-66610-5_15.

133. D. H. Schunk. "Self-Regulated Learning: The Educational Legacy of Paul R. Pintrich". *Educational Psychologist* 40:2, 2005, pp. 85–94. DOI: 10.1207/s15326985ep4002_3.

134. B. Schwendimann, M. Rodríguez-Triana, A. Vozniuk, L. Prieto, M. Shirvani Boroujeni, A. Holzer, D. Gillet, and P. Dillenbourg. "Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research". *IEEE Transactions on Learning Technologies* 10:1, 2017, pp. 30–41. DOI: 10.1109/TLT.2016.2599522.

135. D. Shah. *By The Numbers: MOOCs in 2019*. 2019. URL: https://www.classcentral.com/report/mooc-stats-2019/ (visited on 11/20/2020).

136. D. Shah. *By The Numbers: MOOCs in 2020*. 2020. URL: https://www.classcentral.com/report/mooc-stats-2020/ (visited on 01/14/2021).

137. M. Sharples, C. D. Kloos, Y. Dimitriadis, S. Garlatti, and M. Specht. "Mobile and Accessible Learning for MOOCs". *Journal of Interactive Media in Education* 2015:1, 2015. DOI: http://doi.org/10.5334/jime.ai.

138. S. B. Shum. *UNESCO Policy Brief: Learning Analytics*. Technical report. 2012. URL: https://iite.unesco.org/publications/3214711/.

139. G. Siemens. "Learning Analytics: Envisioning a Research Discipline and a Domain of Practice". In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. LAK '12. Association for Computing Machinery, 2012, pp. 4–8. DOI: 10.1145/2330601.2330605.

140. G. Siemens. "Learning Analytics: The Emergence of a Discipline". *American Behavioral Scientist* 57:10, 2013, pp. 1380–1400. DOI: 10.1177/0002764213498851.

141. G. Siemens. "Massive Open Online Courses: Innovation in Education?" *Open Educational Resources: Innovation, Research and Practice* 5, 2013, pp. 5–16. URL: https://oerknowledgecloud.org/sites/oerknowledgecloud.org/files/pub_PS_OER-IRP_web.pdf#page=31.

142. G. Siemens and R. S. J. d. Baker. "Learning Analytics and Educational Data Mining: Towards Communication and Collaboration". In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*. LAK '12.

143. G. Siemens and D. Gasevic. "Guest Editorial - Learning and Knowledge Analytics". *Journal of Educational Technology & Society* 15:3, 2012, pp. 1–2. URL: http://jstor.org/stable/jeductechsoci.15.3.1.

144. G. Siemens and P. Long. "Penetrating the Fog: Analytics in Learning and Education". *EDUCAUSE Review* 46:5, 2011. URL: https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education.

145. Society for Learning Analytics Research (SoLAR). *What is Learning Analytics?* 2021. URL: https://www.solaresearch.org/about/what-is-learning-analytics/ (visited on 03/16/2021).

146. T. Staubitz and C. Meinel. "Collaboration and Teamwork on a MOOC Platform: A Toolset". In: *Proceedings of the Fourth ACM Conference on Learning @ Scale*. L@S '17. Association for Computing Machinery, 2017, pp. 165–168. DOI: 10.1145/3051457.3053975.

147. T. Staubitz and C. Meinel. "Graded Team Assignments in MOOCs: Effects of Team Composition and Further Factors on Team Dropout Rates and Performance". In: *Proceedings of the Sixth ACM Conference on Learning @ Scale*. L@S '19. Association for Computing Machinery, 2019, pp. 1–10. DOI: 10.1145/3330430.3333619.

148. T. Staubitz, D. Petrick, M. Bauer, J. Renz, and C. Meinel. "Improving the Peer Assessment Experience on MOOC Platforms". In: *Proceedings of the Third ACM Conference on Learning @ Scale*. L@S '16. Association for Computing Machinery, 2016, pp. 389–398. DOI: 10.1145/2876034.2876043.

149. D. E. Steere and D. Cavaiuolo. "Connecting outcomes, goals, and objectives in transition planning". *Teaching Exceptional Children* 34:6, 2002, pp. 54–59. DOI: 10.1177/004005990203400608.

150. K. Stephens-Martinez, M. A. Hearst, and A. Fox. "Monitoring MOOCs: Which Information Sources Do Instructors Value?" In: *Proceedings of the First ACM Conference on Learning @ Scale Conference*. L@S '14. Association for Computing Machinery, 2014, pp. 79–88. DOI: 10.1145/2556325.2566246.

151. D. Suthers and K. Verbert. "Learning Analytics as a "Middle Space"". In: *Proceedings of the Third International Conference on Learning Analytics & Knowledge*. LAK '13. Association for Computing Machinery, 2013, pp. 1–4. DOI: 10.1145/2460296.2460298.

152. Y. Tabaa and A. Medouri. "LASyM: A Learning Analytics System for MOOCs". *International Journal of Advanced Computer Science and Applications* 4:5, 2013. DOI: 10.14569/IJACSA.2013.040516.

153. B. Tabuenca, M. Kalz, H. Drachsler, and M. Specht. "Time will tell: The role of mobile learning analytics in self-regulated learning". *Computers & Education* 89, 2015, pp. 53–74. DOI: 10.1016/j.compedu.2015.08.004.

154. R. Teusner, K. Rollmann, and J. Renz. "Taking Informed Action on Student Activity in MOOCs". In: *Proceedings of the Fourth ACM Conference on Learning @ Scale*. L@S '17. Association for Computing Machinery, 2017, pp. 149–152. DOI: 10.1145/3051457.3053971.

155. The Economist Newspaper Limited. *The world's most valuable resource is no longer oil, but data*. 2017. URL: https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data/ (visited on 03/22/2021).

156. The European Parliament and the Council of the European Union. *Regulation (EU) 2016/679 (General Data Protection Regulation)*. 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 (visited on 06/15/2018).

157. H. Thüs, M. A. Chatti, E. Yalcin, C. Pallasch, B. Kyryliuk, T. Mageramov, and U. Schroeder. "Mobile Learning in Context". *International Journal of Technology Enhanced Learning* 4:5/6, 2012, pp. 332–344. DOI: 10.1504/IJTEL.2012.051818.

158. M. Totschnig, C. Willems, and C. Meinel. "openHPI: Evolution of a MOOC Platform from LMS to SOA". In: *Proceedings of the 5th International Conference on Computer Supported Education*. Vol. 1. CSEDU '13. SciTePress, 2013, pp. 593–598. DOI: 10.5220/0004416905930598.

159. H. Utunen, N. Ndiaye, C. Piroux, R. George, M. Attias, and G. Gamhewage. "Global Reach of an Online COVID-19 Course in Multiple Languages on OpenWHO in the First Quarter of 2020: Analysis of Platform Use Data". *Journal of Medical Internet Research* 22:4, 2020, e19076. DOI: 10.2196/19076.

160. K. Verbert, E. Duval, J. Klerkx, S. Govaerts, and J. Santos. "Learning Analytics Dashboard Applications". *American Behavioral Scientist* 57:10, 2013, pp. 1500–1509. DOI: 10.1177/0002764213479363.

161. K. Verbert, S. Govaerts, E. Duval, J. L. Santos, F. V. Assche, G. Parra, and J. Klerkx. "Learning dashboards: an overview and future research opportunities". *Personal and Ubiquitous Computing* 18, 2014, pp. 1499–1514. DOI: 10.1007/s00779-013-0751-2.

162. O. Viberg, M. Khalil, and M. Baars. "Self-Regulated Learning and Learning Analytics in Online Learning Environments: A Review of Empirical Research". In: *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*. LAK '20. Association for Computing Machinery, 2020, pp. 524–533. DOI: 10.1145/3375462.3375483.

163. L. Vigentini, A. Clayphan, and M. Chitsaz. "Dynamic dashboard for educators and students in FutureLearn MOOCs: experiences and insights". In: *Proceedings of the Joint MOOCs Workshops from the Learning Analytics and Knowledge Conference 2017*. LAK '17. CEUR-WS.org, 2017, pp. 20–35. URL: http://ceur-ws.org/Vol-1967/FLMOOCS_Paper2.pdf.

164. L. Vigentini, A. Clayphan, X. Zhang, and M. Chitsaz. "Overcoming the MOOC Data Deluge with Learning Analytic Dashboards". In: *Learning Analytics: Fundaments, Applications, and Trends: A View of the Current State of the Art to Enhance e-Learning*. Ed. by A. Peña-Ayala. Springer International Publishing, 2017, pp. 171–198. DOI: 10.1007/978-3-319-52977-6_6.

165. M. A. Virtanen, E. Haavisto, E. Liikanen, and M. Kääriäinen. "Ubiquitous learning environments in higher education: A scoping literature review". *Education and Information Technologies* 23, 2018, pp. 985–998. DOI: 10.1007/s10639-017-9646-6.

166. I. de Waard, A. Koutropoulos, N. Keskin, S. C. Abajian, R. Hogue, C. O. Rodriguez, and M. S. Gallagher. "Exploring the MOOC format as a pedagogical approach for mLearning". In: *Proceedings of 10th World Conference on Mobile and Contextual Learning*. mLearn '11. 2011.

167. M. Weller. *Virtual Learning Environments: Using, Choosing and Developing your VLE*. Routledge, 2007.

168. J. Wilkowski, A. Deutsch, and D. M. Russell. "Student Skill and Goal Achievement in the Mapping with Google MOOC". In: *Proceedings of the First ACM Conference on Learning @ Scale Conference*. L@S '14. Association for Computing Machinery, 2014, pp. 3–10. DOI: 10.1145/2556325.2566240.

169. P. H. Winne. "Learning Analytics for Self-Regulated Learning". In: *The Handbook of Learning Analytics*. Ed. by C. Lang, G. Siemens, A. F. Wise, and D. Gaševic. 1st ed. Society for Learning Analytics Research (SoLAR), 2017, pp. 241–249. DOI: 10.18608/hla17.012.

170. M. Yeomans and J. Reich. "Planning Prompts Increase and Forecast Course Completion in Massive Open Online Courses". In: *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*. LAK '17. Association for Computing Machinery, 2017, pp. 464–473. DOI: 10.1145/3027385.3027416.

171. Y. Yoo, H. Lee, I.-H. Jo, and Y. Park. "Educational Dashboards for Smart Learning: Review of Case Studies". In: *Emerging Issues in Smart Learning*. Springer Berlin Heidelberg, 2015, pp. 145–155. DOI: 10.1007/978-3-662-44188-6_21.

172. X. Zhang, C. Li, S. W. Li, and V. Zue. “Automated Segmentation of MOOC Lectures towards Customized Learning”. In: *Proceedings of the IEEE 16th International Conference on Advanced Learning Technologies*. ICALT ’16. IEEE, 2016, pp. 20–22. DOI: 10.1109/ICALT.2016.25.

173. S. Zheng, M. B. Rosson, P. C. Shih, and J. M. Carroll. “Understanding Student Motivation, Behaviors and Perceptions in MOOCs”. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. CSCW ’15. Association for Computing Machinery, 2015, pp. 1882–1895. DOI: 10.1145/2675133.2675217.

174. B. J. Zimmerman. “Attaining Self-Regulation: A Social Cognitive Perspective”. In: *Handbook of Self-Regulation*. Ed. by M. Boekaerts, P. R. Pintrich, and M. Zeidner. Academic Press, 2000, pp. 13–39. DOI: 10.1016/B978-012109890-2/50031-7.

List of Tables

List of Figures

List of Listings

A Analytics Events

Overview of all analytics events on the platform. Deprecated or temporary events are omitted.

Category	Event	Description
Course	enrolled	A user has enrolled for a course.
	un_enrolled	A user has unenrolled from a course.
	completed_course	A user has completed a course.
	share_button_click	A user has clicked on the share button of a course.
	share_open_badge	A user has clicked on the share button of an open badge.
Navigation	visited_item	A user has opened a learning item.
	navigated_next_item	A user has opened the next learning item.
	navigated_prev_item	A user has opened the previous learning item.
	visited_pinboard	A user has opened the course forum.
	visited_question	A user has opened a forum thread.
	visited_progress	A user has opened their course progress.
	visited_peer_assessment_results	A user has opened their peer assessment results.
	visited_announcements	A user has opened the course announcements overview.
	visited_announcement_detail	A user has opened the an announcement.
	visited_learning_rooms	A user has opened the collab spaces overview.
	visited_recap	A user has opened the recap tool.
	visited_profile	A user has opened their account profile.
	visited_dashboard	A user has opened their account dashboard.
	visited_documents	A user has opened their certificates overview.
	visited_downloads	A user has opened their downloads overview.
	visited_preferences	A user has opened their preferences.
	visited_course_dashboard	An admin has opened the course dashboard.
	visited_course_statistics	An admin has opened a course statistics page.
	visited_item_statistics	An admin has opened an item statistics page.
Assignments	submitted_quiz	A user has submitted a quiz.
	submitted_lti(_v2)	A user has submitted an LTI exercise.
Recap	recap_started	A user has started the recap tool.
	recap_stopped	A user has stopped the recap tool.
	recap_answered	A user has submitted answers in the recap tool.

Table continued from previous page.

Category	Event	Description
Video Player	video_play	A user has played a video.
	video_seek	A user has seeked to another position in a video.
	video_pause	A user has paused a video.
	video_end	A user has watched a video to the end.
	video_close	A user has closed a video.
	video_slide_seek	A user has seeked in a video with the slide preview.
	video_transcript_seek	A user has seeked in a video with the transcript.
	video_fullscreen	A user has resized a video to fullscreen.
	video_landscape	A user has rotated a video to landscape (mobile).
	video_portrait	A user has rotated a video to portrait (mobile).
	video_subtitle	A user has toggled the subtitles of a video.
	video_transcript	A user has toggled the transcript of a video.
	video_change_quality	A user has changed the quality of a video.
	video_change_size	A user has changed the size of a video.
	video_change_speed	A user has changed the playback speed of a video.
	video_dual_stream_change	A user has toggled the dual stream mode of a video.
Downloads	downloaded_hd_video	A user has downloaded the HD quality of a video.
	downloaded_sd_video	A user has downloaded the SD quality of a video.
	downloaded_hls_video	A user has downloaded the HLS playlist of a video.
	downloaded_audio	A user has downloaded the audio track of a video.
	downloaded_slides	A user has downloaded the PDF slides of a video.
	downloaded_subtitles	A user has downloaded the subtitles of a video.
	downloaded_reading_material	A user has downloaded the reading material of a video.
	downloaded_section	A user has downloaded video assets for a course section.
	downloaded_certificate	A user has downloaded a course certificate.
	downloaded_open_badge	A user has downloaded an open badge.
Forum	asked_question	A user has opened a forum thread.
	answered_question	A user has answered a forum thread.
	answer_accepted	A user has accepted a forum thread answer.
	commented	A user has commented a forum thread.
	toggled_subscription	A user has subscribed to a forum thread.
Support	helpdesk_opened	A user has opened the helpdesk.
	helpdesk_closed	A user has closed the helpdesk.
	helpdesk_ticket_created	A user has created a support ticket.

B Sample Courses

Overview of all courses that are subject to experiments and evaluations throughout this work, ordered alphabetically by their course code.

bigdata2017: Big Data Analytics

`https://open.hpi.de/courses/bigdata2017/`

Language	German					
Term	6 Weeks from Nov 06, 2017, to Jan 03, 2018					
Enrollments	Start	8,118	Middle	9,373	End	10,093
Shows	-		Middle	5,169	End	6,108
Completion	CoP	2,575	RoA	916	QC	69
Examination	Weekly Assignments and Final Exam					
Study	Learning Behavior of Learners with and without Mobile App Usage (Section 2.3)					
Description	In the course, you will learn how to proceed sensibly in the evaluation of huge amounts of data - starting with the most modern data mining techniques for "digging" previously hidden or unused information, to the preparation and analysis of the data. Current applications and memorable practical examples will familiarize you with the basic problems. Which different algorithms can help to solve them will also be discussed. Finally, we will present common methods that will enable you to evaluate data mining solutions for concrete applications.					

cp5: Cloud-Native Development with SAP Cloud Platform						
https://open.sap.com/courses/cp5/						
Language	English					
Term	6 Weeks from Apr 10, 2018, to May 30, 2018					
Enrollments	Start	5,694	Middle	7,661	End	9,286
Shows	-		Middle	3,619	End	5,249
Completion	CoP	1,560	RoA	627	QC	-
Examination	Weekly Assignments and Final Exam					
Study	Learning Behavior of Learners with and without Mobile App Usage (Section 2.3)					
Description	In this new advanced course, you'll learn how to develop microservice-based cloud-native applications with SAP Cloud Platform through hands-on exercises. You'll be working primarily with the Cloud Foundry environment within SAP Cloud Platform and many of its open-source services to develop step-by-step a Java-based application that is made for the cloud.					

digitalhealth2020: Digital Health for Beginners						
https://open.hpi.de/courses/digitalhealth2020/						
Language	German					
Term	2 Weeks from Sep 22, 2020, to Oct 06, 2020					
Enrollments	Start	4,059	Middle	4,537	End	5,034
Shows	-		Middle	2,228	End	3,467
Completion	CoP	2,431	RoA	1,693	QC	-
Examination	Weekly Assignments					
Study	Completion Rates of the 3 Learner Dashboard Variants and SRL Support (Section 3.4)					
Description	The two-week course provides an up-to-date overview of the topic of digital health. It provides basics, practical exercises and further excursions in the topics of our research. In the course, we also discuss ethical and legal aspects that are extremely important when dealing with health and patient data. Participants will gain insight into the challenges and exciting trends in artificial intelligence, sensor technologies and precision medicine. The aim of the course is to provide a broad overview of the aspects of Digital Health. In addition, the course team wants to create understanding for possible applications and optimization approaches that affect the health of each individual as well as the entire healthcare system.					

imdb2017: In-Memory Data Management

https://open.hpi.de/courses/imdb2017/

Language	English
Term	6 Weeks from Sep 18, 2017, to Nov 18, 2017
Enrollments	Start 4,683 | Middle 5,276 | End 5,825
Shows	- | Middle 2,402 | End 3,128
Completion	CoP 1,078 | RoA 453 | QC 27
Examination	Weekly Assignments and Final Exam
Study	Achievement of Self-Reported Learning Objectives (Section 3.5)
Description	The 'In-Memory Data Management' MOOC in 2017 is the fifth iteration of Prof. Hasso Plattner's successful introduction into the inner mechanics of this recent technology. It is a repetition of the 2015 course and builds on the same, revised material. The course focuses on the management of enterprise data in column-oriented in-memory databases. Latest hardware and software trends led to the development of a new revolutionary database technology that enables flexible and lightning-fast analysis of massive amounts of enterprise data. The basic concepts and design principles of this technology are explained in detail. Beyond that, the implications of the underlying design principles for future enterprise applications and their development are discussed. The MOOC will explain in detail the differences and advantages of an in-memory column-oriented database in contrast to traditional row-oriented disk-based storages.

informationssicherheit2019: Data Security on the Internet

https://open.hpi.de/courses/informationssicherheit2019/

Language	German
Term	2 Weeks from Jan 16, 2019, to Jan 30, 2019
Enrollments	Start 4,354 | Middle 4,947 | End 6,002
Shows	- | Middle 2,967 | End 4,737
Completion	CoP 4,525 | RoA 2,030 | QC -
Examination	Weekly Assignments
Study	Usefulness and Value of the Learner Dashboard (Section 3.3)
Description	In this course, we will shed light on how and whether their connection to online banking is secure or the content of an e-mail is trustworthy. To do this, we will look at the basics of cryptography, security objectives and different types of encryption. In addition, there will be insights into different models and standards used in practice.

international-teams2019: Introduction to Successful Remote Teamwork

https://open.hpi.de/courses/international-teams2019/

Language	English					
Term	4 Weeks from Oct 02, 2019, to Oct 30, 2019					
Enrollments	Start	2,327	Middle	2,778	End	2,991
Shows	-		Middle	1,074	End	1,425
Completion	CoP	370	RoA	212	QC	-
Examination	Team Peer Assessment and Final Exam					
Study	Achievement of Personalized Learning Objectives (Section 3.7)					
Description	Tele-working is becoming a more and more popular topic amongst modern organizations. However, it also comes with some challenges for both: tele-workers and management. This course will make you and your team fit for virtual collaboration in geographically distributed contexts. You will learn about the benefits and risks of driving a virtual team culture and how guided remote work drives to success. Furthermore, you will learn how to use intercultural competences as a key factor of interaction and communication. In the hands-on part of the course you will learn how to select appropriate online collaboration tools and how to employ them in a practical task. Working with a 'real life' virtual team, you will gain first-hand experience about the opportunities and challenges of tele-working.					

internetworking2020: A Half Century of Internet: How it works today

https://open.hpi.de/courses/internetworking2020/

Language	English					
Term	6 Weeks from Sep 01, 2020, to Oct 20, 2020					
Enrollments	Start	3,440	Middle	4,823	End	5,398
Shows	-		Middle	2,131	End	2,827
Completion	CoP	809	RoA	582	QC	17
Examination	Weekly Assignments and Final Exam					
Study	Completion Rates of the 3 Learner Dashboard Variants and SRL Support (Section 3.4)					
Description	The Internet connects more than half of the world's population. This revolutionary form of transmitting all kinds of data between places on the planet has made the network of networks the indispensable backbone of societies. The number of users has exploded to four billion people. The speed of change is dramatic and for some breathtaking. Many well-known and even more unknown personalities have shaped the development of the Internet. However, this exciting success story also reveals the dark sides of this development. What has become of the original hope for a democratization of communication? To what extent has the Internet provided access to better educational opportunities? How do large Internet companies and governments use the Internet? How can you safely communicate over this network?					

intsec2018: Internet Security for Beginners						
`https://open.hpi.de/courses/intsec2018/`						
Language	English					
Term	6 Weeks from Feb 26, 2018, to Apr 27, 2018					
Enrollments	Start	6,242	Middle	7,734	End	8,147
Shows	-		Middle	4,564	End	5,094
Completion	CoP	1,745	RoA	971	QC	46
Examination	Weekly Assignments and Final Exam					
Study	Learning Behavior of Learners with and without Mobile App Usage (Section 2.3)					
Description	Protection from Internet risks requires more action worldwide: from businesses, institutions, public authorities, and every one of us. In this 6-week free course in English, we offer practical support to face this challenging task. With no prior knowledge required, participants can find out what methods hackers use to break into computers and networks. Learn how cybercriminals manage to steal passwords and how you can protect yourself from such cyberattacks.					

javaeinstieg2017: Object-Oriented Programming in Java						
`https://open.hpi.de/courses/javaeinstieg2017/`						
Language	German					
Term	4 Weeks from Mar 27, 2017, to May 14, 2017					
Enrollments	Start	7,127	Middle	9,242	End	10,402
Shows	-		Middle	6,610	End	8,015
Completion	CoP	3,883	RoA	2,124	QC	-
Examination	Programming Exercises, Weekly Assignments and a Bonus Team Peer Assessment					
Study	Achievement of Self-Reported Learning Objectives (Section 3.5)					
Description	In this openHPI beginner's course, we will deal with the basics of object-oriented programming and solve a mysterious kidnapping case together with Detective Duke. An essential feature of object-oriented programming is the appropriate distribution of tasks to components, each of which has its own properties and behaviors and can influence each other. Through regular programming assignments, participants apply what they have learned and acquire practical skills in the Java programming language. The course is rounded off by an in-depth excursus on modeling classes and their dependencies.					

javawork2017: Introduction into a Java IDE						
https://open.hpi.de/courses/javawork2017/						
Language	German					
Term	2 Weeks from May 01, 2017, to May 15, 2017					
Enrollments	Start	3,881	Middle	4,112	End	4,336
Shows	-		Middle	1,482	End	2,096
Completion	CoP	1,481	RoA	194	QC	-
Examination	Peer Assessment					
Study	Achievement of Self-Reported Learning Objectives (Section 3.5)					
Description	This two-week MOOC workshop is offered as a supplement to the 'Object-Oriented Programming in Java 2017' course. Our course participants will be given an introduction to the use of a Java IDE (Integrated Development Environment). With this course we would like to enable our participants to consolidate the knowledge they have learned in the programming course and to implement their own programs.					

kieinstieg2020: Artificial Intelligence and Machine Learning for Beginners						
https://open.hpi.de/courses/kieinstieg2020/						
Language	German					
Term	4 Weeks from Sep 08, 2020, to Oct 06, 2020					
Enrollments	Start	9,268	Middle	10,515	End	11,284
Shows	-		Middle	5,492	End	6,856
Completion	CoP	3,847	RoA	2,775	QC	-
Examination	Weekly Assignments and Final Exam					
Study	Completion Rates of the 3 Learner Dashboard Variants and SRL Support (Section 3.4)					
Description	Here, young people and other interested people without programming experience and technical background knowledge learn to understand the world of machine learning and artificial intelligence. We will introduce you to the basic concepts. You will learn about the differences between traditional programming and the development of self-learning software. Using examples, you will learn what supervised, unsupervised and reinforcement learning are. These concepts form the core for the algorithms that make machine learning work. Experience how such a learning process can be used to recognize patterns and structures in large amounts of data by means of a practical application. Ethical issues in the use of artificial intelligence and the limitations of machine learning technology are also addressed in the free four-week course.					

knowledgegraphs2020: Knowledge Graphs

`https://open.hpi.de/courses/knowledgegraphs2020/`

Language	English					
Term	6 Weeks from Oct 27, 2020, to Dec 15, 2020					
Enrollments	Start	4,812	Middle	6,081	End	6,511
Shows	-		Middle	2,994	End	3,527
Completion	CoP	947	RoA	468	QC	29
Examination	Weekly Assignments and Final Exam					
Study	Completion Rates of the 3 Learner Dashboard Variants and SRL Support (Section 3.4)					
Description	In this course you will learn what is necessary to design, implement, and use knowledge graphs. The focus of this course will be on basic semantic technologies including the principles of knowledge representation and symbolic AI. This includes information encoding via RDF triples, knowledge representation via ontologies with OWL, efficiently querying knowledge graphs via SPARQL, latent representation of knowledge in vector space, as well as knowledge graph applications in innovative information systems, as e.g., semantic and exploratory search.					

learningtheory2020: Computational Learning Theory and Beyond

`https://open.hpi.de/courses/learningtheory2020/`

Language	English					
Term	2 Weeks from Oct 06, 2020, to Oct 27, 2020					
Enrollments	Start	2,312	Middle	2,855	End	3,012
Shows	-		Middle	1,246	End	1,638
Completion	CoP	350	RoA	61	QC	-
Examination	Weekly Assignments and Final Exam					
Study	Completion Rates of the 3 Learner Dashboard Variants and SRL Support (Section 3.4)					
Description	In this course you will be introduced to computational learning theory and get a glimpse of other research towards a theory of artificial intelligence. Our starting point will be a hands-on binary classification task. Basically, this is the challenge of classifying the elements of a given set into two groups (predicting which group each one belongs to) on the basis of given labeled data. Thus the goal of the supervised machine learning algorithms is to derive a correct classification rule. Our interest lies in strategies that work not only for one specific classification task but more universally for a pre-specified set of such. You will get to know a formalization of the aforementioned notions and see illustrating examples. In the main part, you will get to know different learning models which are all based on a modular design. By investigating the learning power of these models and the learnability of the prominent set of half-spaces, we also give arguments for how to choose an appropriate one.					

mainframes2017: Mainframes

`https://open.hpi.de/courses/mainframes2017/`

Language	German					
Term	6 Weeks from Jun 05, 2017, to Jul 27, 2017					
Enrollments	Start	2,635	Middle	3,026	End	3,396
Shows	-		Middle	1,670	End	2,115
Completion	CoP	929	RoA	438	QC	23
Examination	Weekly Assignments and Final Exam					
Study	Achievement of Self-Reported Learning Objectives (Section 3.5)					
Description	Prof. Andreas Polze will lead the course 'Mainframes', which will feature speakers from the Academic Mainframe Consortium. This consortium promotes academic education on mainframe computers and has, among other things, the goal of promoting and advancing research and development in this area. The course is an introduction to the subject and will specifically address mainframe architecture and application development issues. In addition, databases and transaction systems, the topic of security and mainframes, and storage management will be covered in the course. This course is intended for anyone who is interested in the topic of mainframes and has a basic knowledge of computer science and algorithms.					

mainframes2018: Mainframe - Crucial Role in Modern Enterprise Computing

`https://open.hpi.de/courses/mainframes2018/`

Language	English					
Term	6 Weeks from Nov 05, 2018, to Dec 17, 2018					
Enrollments	Start	2,270	Middle	2,534	End	2,614
Shows	-		Middle	974	End	1,096
Completion	CoP	439	RoA	224	QC	18
Examination	Weekly Assignments and Final Exam					
Study	Acceptance and Usefulness of Personalized Learning Objectives (Section 3.6)					
Description	The digital transformation poses a big challenge for enterprises today, and the modern mainframe plays a crucial role in addressing this challenge. As the typical platform for the 'systems of record' of most large organizations, its unique features support today's requirements regarding performance, flexibility and security and enable companies to successfully manage the challenges of an online world. This course will give an overview of the modern mainframe, its concepts and features and their role in enterprise computing. Topics addressed range from mainframe architecture, hardware and operating systems (z/OS, z/VM, Linux), Mainframe application development and transaction processing to state-of-the-art workloads such as blockchain and analytics. In addition, industry success stories will be presented.					

ml2: Enterprise Deep Learning with TensorFlow						
`https://open.sap.com/courses/ml2/`						
Language	English					
Term	6 Weeks from Oct 23, 2017, to Dec 14, 2017					
Enrollments	Start	5,704	Middle	8,568	End	10,212
Shows	-		Middle	5,855	End	7,276
Completion	CoP	3,169	RoA	1,569	QC	-
Examination	Weekly Assignments and Final Exam					
Study	Learning Behavior of Learners with and without Mobile App Usage (Section 2.3)					
Description	The objective of this course is to provide a hands-on introduction to deep learning, with emphasis on practical enterprise applications. Taking an engineering approach to deep learning, the course focuses on building deep neural network models for typical enterprise problems, including when to use deep learning, examples of industry applications, and how to deploy deep learning in enterprise systems. The course features experts from academia and industry to show different perspectives on deep learning. All examples are implemented using the TensorFlow deep learning framework.					

prototype2019: Human-Centered Design: Building and Testing Prototypes						
`https://open.hpi.de/courses/prototype2019/`						
Language	English					
Term	4 Weeks from Aug 28, 2019, to Oct 10, 2019					
Enrollments	Start	3,029	Middle	3,356	End	3,533
Shows	-		Middle	1,568	End	1,851
Completion	CoP	626	RoA	250	QC	-
Examination	Three Exercises and Peer Assessment					
Study	Achievement of Personalized Learning Objectives (Section 3.7)					
Description	This course introduces you to helpful skills for making an idea tangible and testing it with potential users. We take a task-based approach to build these skills. You will learn to take an idea to the next level and build a simple prototype, plan a testing scenario and collect feedback. This MOOC builds on the 2018 ‘From Synthesis to Creative Ideas’ course, but you can also take it as a stand-alone MOOC. The assignments are designed for individual work.					

s4h12: Intelligent ERP with SAP S/4HANA Cloud

https://open.sap.com/courses/s4h12/

Language	English					
Term	3 Weeks from Nov 07, 2018, to Dec 06, 2018					
Enrollments	Start	13,512	Middle	15,258	End	16,539
Shows	-		Middle	5,142	End	6,981
Completion	CoP	4,248	RoA	1,786	QC	-
Examination	Weekly Assignments and Final Exam					
Study	Acceptance and Usefulness of Personalized Learning Objectives (Section 3.6)					
Description	Join this free open online course to learn how you can benefit from the automation of business processes using machine learning, prepare for the future using predictive analytics, and operate your intelligent cloud ERP hands-free based on natural language interaction. In the first week of this course, we'll set the foundation and provide information on our key pillars for intelligence, such as automation primarily achieved by machine learning, and digital user experience. After setting the scene, we'll present several use cases across LoBs that showcase the built-in intelligence in SAP S/4HANA Cloud. In the second week, you'll hear how our intelligent technologies, such as machine learning, predictive analytics, and natural language interaction, can be used in practice in the areas of finance, procurement and EPPM. In the third week, we'll conclude the use cases with the areas of sales and manufacturing. Completing the course, you'll learn in detail how innovative technologies such as machine learning, predictive analytics, and digital user experience using SAP CoPilot can be utilized in practice to gain an intelligence edge over the competition.					

searchengine2017: How does a search engine work?

https://open.hpi.de/courses/searchengine2017/

Language	German					
Term	2 Weeks from May 29, 2017, to Jun 20, 2017					
Enrollments	Start	3,922	Middle	4,145	End	4,484
Shows	-		Middle	1,702	End	2,660
Completion	CoP	2,081	RoA	814	QC	-
Examination	Final Exam					
Study	Achievement of Self-Reported Learning Objectives (Section 3.5)					
Description	In this course we want to introduce the technical basics and take a closer look at simple concepts of information retrieval in the web context: How is a search engine built? What happens when I type in a search query? What criteria are used to create lists of results? Topics such as search engine optimization (SEO) or the specific use of certain search engines will not be discussed in detail.					

C Study Data

C.1 Usefulness and Value of the Learner Dashboard

Survey data analyzed in Section 3.3. Number of Participants: 217.

Q1:	**The learning dashboard is easy to use.**
Q2:	**The functions of the learning dashboard are exactly right for my goals.**
Q3:	**It is quickly apparent how to use the learning dashboard.**
Q4:	**I consider the learning dashboard extremely useful.**
Q5:	**The operating procedures of the learning dashboard are simple to understand.**
Q6:	**With the help of the learning dashboard I will achieve my goals.**

	strongly disagree -2	disagree -1	0	agree +1	strongly agree +2	no answer
Q1	1 / 0.46%	1 / 0.46%	7 / 3.23%	87 / 40.09%	118 / 54.38%	3 / 1.38%
Q2	2 / 0.92%	6 / 2.76%	18 / 8.29%	113 / 52.07%	74 / 34.10%	4 / 1.84%
Q3	1 / 0.46%	5 / 2.30%	16 / 7.37%	91 / 41.94%	101 / 46.54%	3 / 1.38%
Q4	3 / 1.38%	6 / 2.76%	21 / 9.68%	86 / 39.63%	98 / 45.16%	3 / 1.38%
Q5	1 / 0.46%	3 / 1.38%	10 / 4.61%	103 / 47.47%	98 / 45.16%	2 / 0.92%
Q6	4 / 1.84%	9 / 4.15%	46 / 21.20%	96 / 44.24%	53 / 24.42%	9 / 4.15%

C1:	**I consider the learning progress overview as useful.**
C2:	**I consider the time estimation for available course material as useful.**
C3:	**I consider the overview of the next course dates as useful.**
C4:	**I consider the overview of my forum activity as useful.**
C5:	**I consider the overview of my self-tests and assignments as useful.**
C6:	**I consider the repetition suggestions as useful.**
C7:	**I consider the overview of my achieved points as useful.**
C8:	**I consider the information on my time spent on quizzes as useful.**
C9:	**I consider the information on the timeliness of my submissions as useful.**
C10:	**I consider the overview of my visited items as useful.**
C11:	**I consider the overview of my learning activities per week as useful.**
C12:	**I consider the overview of my learning sessions as useful.**

	strongly disagree -2	disagree -1	0	agree +1	strongly agree +2	no answer
C1	0 / 0.00%	5 / 2.30%	6 / 2.76%	73 / 33.64%	132 / 60.83%	1 / 0.46%
C2	8 / 3.69%	8 / 3.69%	27 / 12.44%	100 / 46.08%	71 / 32.72%	3 / 1.38%
C3	2 / 0.92%	4 / 1.84%	13 / 5.99%	98 / 45.16%	97 / 44.70%	3 / 1.38%
C4	6 / 2.76%	13 / 5.99%	78 / 35.94%	77 / 35.48%	25 / 11.52%	18 / 8.29%
C5	2 / 0.92%	3 / 1.38%	12 / 5.53%	108 / 49.77%	87 / 40.09%	5 / 2.30%
C6	3 / 1.38%	6 / 2.76%	23 / 10.60%	101 / 46.54%	74 / 34.10%	10 / 4.61%
C7	0 / 0.00%	6 / 2.76%	14 / 6.45%	100 / 46.08%	93 / 42.86%	4 / 1.84%
C8	5 / 2.30%	11 / 5.07%	47 / 21.66%	96 / 44.24%	54 / 24.88%	4 / 1.84%
C9	7 / 3.23%	8 / 3.69%	44 / 20.28%	98 / 45.16%	53 / 24.42%	7 / 3.23%
C10	7 / 3.23%	12 / 5.53%	39 / 17.97%	92 / 42.40%	56 / 25.81%	11 / 5.07%
C11	10 / 4.61%	11 / 5.07%	62 / 28.57%	83 / 38.25%	42 / 19.35%	9 / 4.15%
C12	9 / 4.15%	16 / 7.37%	64 / 29.49%	80 / 36.87%	38 / 17.51%	10 / 4.61%

How did you like the new learning dashboard in general?

1⋆	2⋆	3⋆	4⋆	5⋆
5 / 2.30%	5 / 2.30%	18 / 8.29%	86 / 39.63%	103 / 47.47%

C.2 Supporting Self-Regulated Learning with the Learner Dashboard

Survey data analyzed in Section 3.4. Number of Participants: 296.

D1: **For the new Learner Dashboard it is clear what data is being collected.**
D2: **For the new Learner Dashboard it is clear why the data is being collected.**
AR1: **The new Learner Dashboard makes me aware of my current learning situation.**
AR2: **The new Learner Dashboard makes me forecast my possible future learning situation given my (un)changed behavior.**
AR3: **The new Learner Dashboard stimulates me to reflect on my past learning behavior.**
AR4: **The new Learner Dashboard stimulates me to adapt my learning behavior if necessary.**
I1: **The new Learner Dashboard stimulates me to study more efficiently.**
I2: **The new Learner Dashboard stimulates me to study more effectively.**

Q	strongly disagree 1	2	3	4	5	6	7	8	9	strongly agree 10
D1	12 / 4.05%	6 / 2.03%	13 / 4.39%	10 / 3.38%	11 / 3.72%	15 / 5.07%	25 / 8.45%	62 / 20.95%	62 / 20.95%	80 / 27.03%
D2	12 / 4.05%	6 / 2.03%	13 / 4.39%	7 / 2.36%	21 / 7.09%	18 / 6.08%	29 / 9.80%	59 / 19.93%	56 / 18.92%	75 / 25.34%
AR1	10 / 3.38%	8 / 2.70%	3 / 1.01%	7 / 2.36%	11 / 3.72%	11 / 3.72%	22 / 7.43%	54 / 18.24%	57 / 19.26%	113 / 38.18%
AR2	19 / 6.42%	10 / 3.38%	14 / 4.73%	13 / 4.39%	29 / 9.80%	25 / 8.45%	42 / 14.19%	66 / 22.30%	38 / 12.84%	40 / 13.51%
AR3	15 / 5.07%	7 / 2.36%	15 / 5.07%	15 / 5.07%	17 / 5.74%	25 / 8.45%	41 / 13.85%	50 / 16.89%	60 / 20.27%	51 / 17.23%
AR4	17 / 5.74%	13 / 4.39%	18 / 6.08%	10 / 3.38%	26 / 8.78%	26 / 8.78%	31 / 10.47%	62 / 20.95%	45 / 15.20%	48 / 16.22%
I1	21 / 7.09%	15 / 5.07%	26 / 8.78%	14 / 4.73%	29 / 9.80%	40 / 13.51%	40 / 13.51%	44 / 14.86%	35 / 11.82%	32 / 10.81%
I2	19 / 6.42%	12 / 4.05%	20 / 6.76%	20 / 6.76%	28 / 9.46%	31 / 10.47%	38 / 12.84%	51 / 17.23%	42 / 14.19%	35 / 11.82%

C.3 Usefulness of Personalized Learning Objectives

Survey data analyzed in Section 3.6. Number of Participants: 163.

Q1: **Did you select a learning objective for the course?**		
Answer	Count	Percent
Yes	126	77.30%
No	21	12.88%
No answer	16	9.82%

Q2: **Did you like the number of available learning objectives?**		
Answer	Count	Percent
I would like to have more learning objectives available.	34	20.86%
The number of learning objectives available was sufficient.	90	55.21%
There were too many learning objectives available.	10	6.13%
No answer	29	17.79%

Q3: **What was your motivation to select a learning objective?**		
Answer (Multi Select)	Count	Percent
I wanted to support the research.	45	35.71%
I wanted to try the (potential) new feature.	72	57.14%
I was interested in the experience of choosing and following an objective for the course.	68	53.97%
I always decide on my personal objective anyway and could now select it on the platform.	35	27.78%
Other	1	0.79%

Q4.1:	**The learning objectives selection is easy to use.**
Q4.2:	**The functions of the learning objectives selection are exactly right for my goals.**
Q4.3:	**It is quickly apparent how to use the learning objectives selection.**
Q4.4:	**I consider the learning objectives selection extremely useful.**
Q4.5:	**The operating procedures of the learning objectives selection are simple to understand.**
Q4.6:	**With the help of the learning objectives selection I will achieve my goals.**
Q4.7:	**I liked the presentation of the learning objectives selection (i. e., using a modal).**

	strongly disagree -2	disagree -1	0	agree +1	strongly agree +2	no answer
Q4.1	0 / 0.00%	3 / 1.84%	10 / 6.13%	51 / 31.29%	61 / 37.42%	38 / 23.31%
Q4.2	2 / 1.23%	7 / 4.29%	14 / 8.59%	62 / 38.04%	33 / 20.25%	45 / 27.61%
Q4.3	1 / 0.61%	6 / 3.68%	9 / 5.52%	60 / 36.81%	47 / 28.83%	40 / 24.54%
Q4.4	4 / 2.45%	5 / 3.07%	24 / 14.72%	53 / 32.52%	36 / 22.09%	41 / 25.15%
Q4.5	1 / 0.61%	4 / 2.45%	12 / 7.36%	56 / 34.36%	49 / 30.06%	41 / 25.15%
Q4.6	2 / 1.23%	8 / 4.91%	27 / 16.56%	52 / 31.90%	31 / 19.02%	43 / 26.38%
Q4.7	4 / 2.45%	1 / 0.61%	18 / 11.04%	55 / 33.74%	42 / 25.77%	43 / 26.38%

Q5.1:	**The information displayed on each learning objective was useful.**
Q5.2:	**The information displayed on each learning objective was sufficient to decide on a particular objective.**
Q5.3:	**I decided on a particular objective based on the provided details (i. e., time effort) rather than the topic covered.**
Q5.4:	**The selection of a learning objective helped me to become clear about my primary interest for the course.**
Q5.5:	**The learning objective helped me to focus on my primary interest for the course.**
Q5.6:	**The learning objective motivated me to commit to the course.**
Q5.7:	**The learning objective motivated me to complete at least the learning material included in the objective.**
Q5.8:	**The learning objective motivated me to complete more than the material I initially intended to complete.**
Q5.9:	**Working towards a selected objective did improve my learning effectiveness.**
Q5.10:	**I achieved my personal learning goals in this MOOC.**

	strongly disagree -2	disagree -1	0	agree +1	strongly agree +2	no answer
Q5.1	2 / 1.23%	0 / 0.00%	13 / 7.98%	67 / 41.10%	42 / 25.77%	39 / 23.93%
Q5.2	1 / 0.61%	4 / 2.45%	18 / 11.04%	71 / 43.56%	29 / 17.79%	40 / 24.54%
Q5.3	9 / 5.52%	16 / 9.82%	23 / 14.11%	51 / 31.29%	23 / 14.11%	41 / 25.15%
Q5.4	3 / 1.84%	8 / 4.91%	27 / 16.56%	54 / 33.13%	29 / 17.79%	42 / 25.77%
Q5.5	2 / 1.23%	10 / 6.13%	20 / 12.27%	59 / 36.20%	28 / 17.18%	44 / 26.99%
Q5.6	3 / 1.84%	13 / 7.98%	22 / 13.50%	52 / 31.90%	28 / 17.18%	45 / 27.61%
Q5.7	4 / 2.45%	11 / 6.75%	19 / 11.66%	54 / 33.13%	30 / 18.40%	45 / 27.61%
Q5.8	4 / 2.45%	13 / 7.98%	31 / 19.02%	45 / 27.61%	25 / 15.34%	45 / 27.61%
Q5.9	6 / 3.68%	9 / 5.52%	23 / 14.11%	52 / 31.90%	30 / 18.40%	43 / 26.38%
Q5.10	3 / 1.84%	4 / 2.45%	14 / 8.59%	56 / 34.36%	43 / 26.38%	43 / 26.38%

Q6: **Did you use the estimated time effort as criterium for selecting a learning objective?**

Answer	Count	Percent
Yes	54	33.33%
No, but I found the information helpful.	62	38.27%
No, because the information was not helpful.	4	2.47%
No answer	43	26.54%

Q7: **To what extent did you focus on the learning resources that were part of your learning objective?**

Answer	Count	Percent
I have only focused on the highlighted content.	31	19.38%
The given learning resources motivated me to work on further learning resources which were not part of the objective.	41	25.62%
I have adhered to the course structure and worked on all learning resources, regardless of my selection.	43	26.88%
No answer	48	29.45%

Q8.1: **The highlighting of the learning resource was helpful.**
Q8.2: **It was clear which learning resources were part of my learning objective.**
Q8.3: **It was clear which sections contain learning resources being part of my learning objective.**

	strongly disagree -2	disagree -1	0	agree +1	strongly agree +2	no answer
Q8.1	4 / 2.45%	2 / 1.23%	10 / 6.13%	70 / 42.94%	36 / 22.09%	41 / 25.15%
Q8.2	2 / 1.23%	2 / 1.23%	11 / 6.75%	71 / 43.56%	36 / 22.09%	41 / 25.15%
Q8.3	1 / 0.61%	2 / 1.23%	12 / 7.36%	70 / 42.94%	35 / 21.47%	43 / 26.38%

Q9: **Did you change your objective during course run time?**

Answer	Count	Percent
Yes	13	7.98%
No	108	66.26%
No answer	42	25.77%

Q10: **Would you like to have learning objectives available in other courses, too?**

Answer	Count	Percent
Yes	105	64.42%
No	18	11.04%
Uncertain	21	12.88%
No answer	19	11.66%

Q11: **Please rate the overall concept of personalized learning objectives.**

1⋆	2⋆	3⋆	4⋆	5⋆	no answer
5 / 3.07%	3 / 1.84%	21 / 12.88%	46 / 28.22%	52 / 31.90%	36 / 22.09%

C.4 Usefulness of Web Analytics Insights and Usability of the Teacher Dashboard

Survey data analyzed in Section 4.2. Number of Participants: 11.

Q01: **The dashboard helps me to monitor the activity of my courses.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	1 / 9.09%	7 / 63.64%	3 / 27.27%
Revised dashboard	0 / 0.00%	1 / 9.09%	6 / 54.55%	4 / 36.36%

Q02: **The dashboard facilitates access to relevant metrics.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	2 / 18.18%	6 / 54.55%	3 / 27.27%
Revised dashboard	0 / 0.00%	2 / 18.18%	2 / 18.18%	7 / 63.64%

Q03: **The dashboard meets my needs.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	5 / 45.45%	5 / 45.45%	1 / 9.09%
Revised dashboard	0 / 0.00%	3 / 27.27%	6 / 54.55%	2 / 18.18%

Q04: **I regularly use the dashboard for my work.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	1 / 9.09%	4 / 36.36%	6 / 54.55%
Revised dashboard	0 / 0.00%	0 / 0.00%	5 / 45.45%	6 / 54.55%

Q05: **The dashboard is easy to use.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	1 / 9.09%	2 / 18.18%	5 / 45.45%	3 / 27.27%
Revised dashboard	0 / 0.00%	2 / 18.18%	3 / 27.27%	6 / 54.55%

Q06: **The dashboard is understandable.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	1 / 9.09%	5 / 45.45%	5 / 45.45%	0 / 0.00%
Revised dashboard	0 / 0.00%	3 / 27.27%	3 / 27.27%	5 / 45.45%

Q07: **The dashboard loads fast.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	1 / 9.09%	8 / 72.73%	2 / 18.18%	0 / 0.00%
Revised dashboard	0 / 0.00%	3 / 27.27%	5 / 45.45%	3 / 27.27%

Q08: **The dashboard is clear and tidy.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	1 / 9.09%	6 / 54.55%	4 / 36.36%	0 / 0.00%
Revised dashboard	0 / 0.00%	2 / 18.18%	3 / 27.27%	6 / 54.55%

Q09: **The dashboard works the way I would expect.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	3 / 27.27%	7 / 63.64%	1 / 9.09%
Revised dashboard	0 / 0.00%	2 / 18.18%	5 / 45.45%	4 / 36.36%

Q10: **I like to use the dashboard.**				
Dashboard version	Strongly disagree (0)	Somewhat disagree (1)	Somewhat agree (2)	Strongly agree (3)
Existing dashboard	0 / 0.00%	4 / 36.36%	7 / 63.64%	0 / 0.00%
Revised dashboard	0 / 0.00%	2 / 18.18%	4 / 36.36%	5 / 45.45%

C.5 Usage of the Teacher Dashboard and Statistics

Platform data analyzed in Section 4.3.

Course Nr.	Platform	Page Visits	By Teachers	Visit Ratio in Course Phases		
				Design and Dev.	Facilitation	Evaluation
1	openHPI	259	11	0.2432	0.6100	0.1467
2	openHPI	233	10	0.0129	0.7639	0.2232
3	openHPI	537	11	0.3464	0.4227	0.2309
4	openHPI	312	8	0.1410	0.5737	0.2853
5	openHPI	687	13	0.1223	0.7496	0.1281
6	openSAP	221	8	0.3575	0.4887	0.1538
7	openSAP	149	5	0.0336	0.7785	0.1879
8	openSAP	49	4	0.0408	0.4082	0.5510
9	openSAP	41	6	0.0000	0.7805	0.2195
10	openSAP	17	4	0.0000	0.4118	0.5882
11	openSAP	118	8	0.0254	0.5424	0.4322
12	openSAP	36	7	0.2778	0.2500	0.4722
13	openSAP	187	5	0.0267	0.5882	0.3850
14	openSAP	16	4	0.0000	0.2500	0.7500
15	openSAP	69	11	0.1884	0.5362	0.2754
16	openSAP	30	9	0.0333	0.3667	0.6000
17	openSAP	82	6	0.0244	0.6707	0.3049
18	openSAP	28	5	0.0000	0.7143	0.2857
19	openSAP	8	3	0.0000	0.2500	0.7500
20	openSAP	19	5	0.0000	0.2632	0.7368
21	openSAP	65	5	0.0154	0.5538	0.4308
22	openSAP	133	8	0.0000	0.4737	0.5263
23	openSAP	132	8	0.4848	0.2879	0.2273
24	openSAP	47	8	0.0426	0.7021	0.2553
25	openSAP	34	6	0.0294	0.4118	0.5588
26	openSAP	22	6	0.1364	0.2273	0.6364
27	openSAP	77	5	0.0519	0.7273	0.2208
28	openSAP	133	7	0.0451	0.7218	0.2331
29	openSAP	32	5	0.0625	0.5938	0.3438
30	openSAP	27	4	0.0370	0.5185	0.4444
31	mooc.house	570	15	0.3737	0.4737	0.1526
32	mooc.house	48	6	0.2917	0.4375	0.2708
33	mooc.house	43	6	0.3488	0.4419	0.2093

C.6 Perception of the Teacher Dashboard and Statistics

Survey data analyzed in Section 4.3. Number of Participants: 23.

How useful do you find the Enrollments and Certificates Overview in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	11 / 47.83%	4 / 17.39%	2 / 8.70%	0 / 0.00%	3 / 13.04%	2 / 8.70%	1 / 4.35%
Design and Dev.	8 / 34.78%	3 / 13.04%	3 / 13.04%	1 / 4.35%	5 / 21.74%	2 / 8.70%	1 / 4.35%
Facilitation	1 / 4.35%	0 / 0.00%	1 / 4.35%	2 / 8.70%	9 / 39.13%	4 / 17.39%	6 / 26.09%
Evaluation	1 / 4.35%	0 / 0.00%	0 / 0.00%	3 / 13.04%	3 / 13.04%	3 / 13.04%	13 / 56.52%

How useful do you find the Learning Items KPIs in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	4 / 17.39%	2 / 8.70%	1 / 4.35%	1 / 4.35%	2 / 8.70%	1 / 4.35%
Design and Dev.	12 / 52.17%	3 / 13.04%	2 / 8.70%	1 / 4.35%	2 / 8.70%	2 / 8.70%	1 / 4.35%
Facilitation	0 / 0.00%	2 / 8.70%	1 / 4.35%	2 / 8.70%	5 / 21.74%	4 / 17.39%	9 / 39.13%
Evaluation	0 / 0.00%	2 / 8.70%	2 / 8.70%	2 / 8.70%	3 / 13.04%	7 / 30.43%	7 / 30.43%

How useful do you find the Forum KPIs in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	13 / 56.52%	5 / 21.74%	2 / 8.70%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Design and Dev.	12 / 52.17%	4 / 17.39%	4 / 17.39%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Facilitation	2 / 8.70%	1 / 4.35%	2 / 8.70%	2 / 8.70%	3 / 13.04%	3 / 13.04%	10 / 43.48%
Evaluation	2 / 8.70%	1 / 4.35%	1 / 4.35%	0 / 0.00%	6 / 26.09%	5 / 21.74%	8 / 34.78%

How useful do you find the Enrollments and Learners Historic Data in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	9 / 39.13%	5 / 21.74%	1 / 4.35%	3 / 13.04%	1 / 4.35%	4 / 17.39%	0 / 0.00%
Design and Dev.	6 / 26.09%	4 / 17.39%	0 / 0.00%	5 / 21.74%	4 / 17.39%	4 / 17.39%	0 / 0.00%
Facilitation	2 / 8.70%	0 / 0.00%	2 / 8.70%	5 / 21.74%	2 / 8.70%	6 / 26.09%	6 / 26.09%
Evaluation	1 / 4.35%	1 / 4.35%	1 / 4.35%	1 / 4.35%	4 / 17.39%	7 / 30.43%	8 / 34.78%

How useful do you find the Forum and Helpdesk Historic Data in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	13 / 56.52%	3 / 13.04%	4 / 17.39%	1 / 4.35%	2 / 8.70%	0 / 0.00%	0 / 0.00%
Design and Dev.	12 / 52.17%	2 / 8.70%	4 / 17.39%	3 / 13.04%	1 / 4.35%	1 / 4.35%	0 / 0.00%
Facilitation	2 / 8.70%	2 / 8.70%	3 / 13.04%	5 / 21.74%	6 / 26.09%	1 / 4.35%	4 / 17.39%
Evaluation	2 / 8.70%	1 / 4.35%	6 / 26.09%	3 / 13.04%	3 / 13.04%	2 / 8.70%	6 / 26.09%

How useful do you find the User Locations in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	5 / 21.74%	1 / 4.35%	0 / 0.00%	1 / 4.35%	1 / 4.35%	3 / 13.04%
Design and Dev.	9 / 39.13%	6 / 26.09%	0 / 0.00%	2 / 8.70%	2 / 8.70%	1 / 4.35%	3 / 13.04%
Facilitation	3 / 13.04%	1 / 4.35%	1 / 4.35%	6 / 26.09%	4 / 17.39%	3 / 13.04%	5 / 21.74%
Evaluation	1 / 4.35%	2 / 8.70%	2 / 8.70%	1 / 4.35%	5 / 21.74%	6 / 26.09%	6 / 26.09%

How useful do you find the Active Learners by Time of Day in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	15 / 65.22%	5 / 21.74%	0 / 0.00%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Design and Dev.	15 / 65.22%	4 / 17.39%	0 / 0.00%	0 / 0.00%	1 / 4.35%	2 / 8.70%	1 / 4.35%
Facilitation	8 / 34.78%	1 / 4.35%	2 / 8.70%	0 / 0.00%	3 / 13.04%	3 / 13.04%	6 / 26.09%
Evaluation	5 / 21.74%	2 / 8.70%	4 / 17.39%	2 / 8.70%	3 / 13.04%	3 / 13.04%	4 / 17.39%

How useful do you find the Age Distribution in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	10 / 43.48%	5 / 21.74%	3 / 13.04%	1 / 4.35%	2 / 8.70%	1 / 4.35%	1 / 4.35%
Design and Dev.	9 / 39.13%	5 / 21.74%	0 / 0.00%	4 / 17.39%	2 / 8.70%	2 / 8.70%	1 / 4.35%
Facilitation	1 / 4.35%	2 / 8.70%	6 / 26.09%	3 / 13.04%	4 / 17.39%	2 / 8.70%	5 / 21.74%
Evaluation	1 / 4.35%	1 / 4.35%	0 / 0.00%	4 / 17.39%	8 / 34.78%	4 / 17.39%	5 / 21.74%

How useful do you find the Client Usage in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	7 / 30.43%	0 / 0.00%	2 / 8.70%	0 / 0.00%	2 / 8.70%	0 / 0.00%
Design and Dev.	11 / 47.83%	4 / 17.39%	2 / 8.70%	3 / 13.04%	0 / 0.00%	2 / 8.70%	1 / 4.35%
Facilitation	5 / 21.74%	1 / 4.35%	4 / 17.39%	5 / 21.74%	0 / 0.00%	3 / 13.04%	5 / 21.74%
Evaluation	2 / 8.70%	2 / 8.70%	2 / 8.70%	5 / 21.74%	2 / 8.70%	5 / 21.74%	5 / 21.74%

How useful do you find the Item Visits Statistics in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	11 / 47.83%	4 / 17.39%	3 / 13.04%	2 / 8.70%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Design and Dev.	11 / 47.83%	4 / 17.39%	2 / 8.70%	2 / 8.70%	2 / 8.70%	1 / 4.35%	1 / 4.35%
Facilitation	1 / 4.35%	2 / 8.70%	2 / 8.70%	5 / 21.74%	5 / 21.74%	2 / 8.70%	6 / 26.09%
Evaluation	0 / 0.00%	1 / 4.35%	0 / 0.00%	4 / 17.39%	8 / 34.78%	3 / 13.04%	7 / 30.43%

How useful do you find the Videos Statistics in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	5 / 21.74%	3 / 13.04%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Design and Dev.	12 / 52.17%	5 / 21.74%	2 / 8.70%	1 / 4.35%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Facilitation	2 / 8.70%	3 / 13.04%	1 / 4.35%	1 / 4.35%	5 / 21.74%	6 / 26.09%	5 / 21.74%
Evaluation	2 / 8.70%	1 / 4.35%	1 / 4.35%	2 / 8.70%	4 / 17.39%	5 / 21.74%	8 / 34.78%

How useful do you find the Downloads Statistics in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	6 / 26.09%	1 / 4.35%	2 / 8.70%	0 / 0.00%	1 / 4.35%	1 / 4.35%
Design and Dev.	12 / 52.17%	5 / 21.74%	2 / 8.70%	2 / 8.70%	0 / 0.00%	1 / 4.35%	1 / 4.35%
Facilitation	3 / 13.04%	4 / 17.39%	1 / 4.35%	1 / 4.35%	6 / 26.09%	3 / 13.04%	5 / 21.74%
Evaluation	4 / 17.39%	1 / 4.35%	1 / 4.35%	1 / 4.35%	3 / 13.04%	3 / 13.04%	10 / 43.48%

How useful do you find the Rich Texts Statistics in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	15 / 65.22%	3 / 13.04%	0 / 0.00%	3 / 13.04%	1 / 4.35%	0 / 0.00%	1 / 4.35%
Design and Dev.	14 / 60.87%	4 / 17.39%	1 / 4.35%	2 / 8.70%	1 / 4.35%	0 / 0.00%	1 / 4.35%
Facilitation	5 / 21.74%	4 / 17.39%	3 / 13.04%	1 / 4.35%	3 / 13.04%	2 / 8.70%	5 / 21.74%
Evaluation	4 / 17.39%	4 / 17.39%	2 / 8.70%	2 / 8.70%	3 / 13.04%	0 / 0.00%	8 / 34.78%

How useful do you find the Quizzes Statistics in the different course phases?

Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	3 / 13.04%	2 / 8.70%	1 / 4.35%	0 / 0.00%	3 / 13.04%	2 / 8.70%
Design and Dev.	10 / 43.48%	2 / 8.70%	4 / 17.39%	2 / 8.70%	0 / 0.00%	4 / 17.39%	1 / 4.35%
Facilitation	1 / 4.35%	0 / 0.00%	4 / 17.39%	2 / 8.70%	3 / 13.04%	7 / 30.43%	6 / 26.09%
Evaluation	0 / 0.00%	0 / 0.00%	2 / 8.70%	0 / 0.00%	6 / 26.09%	5 / 21.74%	10 / 43.48%

How useful do you find the Rich Text Item Details Statistics in the different course phases?							
Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	13 / 56.52%	4 / 17.39%	1 / 4.35%	3 / 13.04%	1 / 4.35%	0 / 0.00%	1 / 4.35%
Design and Dev.	12 / 52.17%	3 / 13.04%	2 / 8.70%	3 / 13.04%	2 / 8.70%	0 / 0.00%	1 / 4.35%
Facilitation	4 / 17.39%	3 / 13.04%	3 / 13.04%	5 / 21.74%	3 / 13.04%	1 / 4.35%	4 / 17.39%
Evaluation	2 / 8.70%	3 / 13.04%	1 / 4.35%	2 / 8.70%	6 / 26.09%	1 / 4.35%	8 / 34.78%

How useful do you find the Video Item Details Statistics in the different course phases?							
Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	14 / 60.87%	4 / 17.39%	1 / 4.35%	1 / 4.35%	0 / 0.00%	2 / 8.70%	1 / 4.35%
Design and Dev.	13 / 56.52%	3 / 13.04%	1 / 4.35%	3 / 13.04%	0 / 0.00%	2 / 8.70%	1 / 4.35%
Facilitation	4 / 17.39%	2 / 8.70%	6 / 26.09%	1 / 4.35%	3 / 13.04%	4 / 17.39%	3 / 13.04%
Evaluation	5 / 21.74%	3 / 13.04%	2 / 8.70%	1 / 4.35%	4 / 17.39%	3 / 13.04%	5 / 21.74%

How useful do you find the Quiz Item Details Statistics in the different course phases?							
Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	12 / 52.17%	3 / 13.04%	2 / 8.70%	2 / 8.70%	0 / 0.00%	1 / 4.35%	3 / 13.04%
Design and Dev.	11 / 47.83%	2 / 8.70%	2 / 8.70%	3 / 13.04%	0 / 0.00%	2 / 8.70%	3 / 13.04%
Facilitation	1 / 4.35%	0 / 0.00%	4 / 17.39%	2 / 8.70%	0 / 0.00%	5 / 21.74%	11 / 47.83%
Evaluation	1 / 4.35%	1 / 4.35%	2 / 8.70%	0 / 0.00%	5 / 21.74%	4 / 17.39%	10 / 43.48%

How useful do you find the Activity Statistics in the different course phases?							
Phase	not useful at all 1	2	3	4	5	6	very useful 7
Pre-Design	14 / 60.87%	5 / 21.74%	1 / 4.35%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Design and Dev.	13 / 56.52%	5 / 21.74%	2 / 8.70%	0 / 0.00%	1 / 4.35%	1 / 4.35%	1 / 4.35%
Facilitation	3 / 13.04%	2 / 8.70%	4 / 17.39%	3 / 13.04%	3 / 13.04%	4 / 17.39%	4 / 17.39%
Evaluation	3 / 13.04%	3 / 13.04%	2 / 8.70%	3 / 13.04%	4 / 17.39%	3 / 13.04%	5 / 21.74%

D1: **For the Course Dashboard and Statistics it is clear what data is being collected.**
D2: **For the Course Dashboard and Statistics it is clear why the data is being collected.**
AR1: **The Course Dashboard and Statistics make me aware of my students' current learning situation.**
AR2: **The Course Dashboard and Statistics make me forecast my students' possible future learning situation given their (un)changed behavior.**
AR3: **The Course Dashboard and Statistics stimulate me to reflect on my past teaching behavior.**
AR4: **The Course Dashboard and Statistics stimulate me to adapt my teaching behavior if necessary.**
I1: **The Course Dashboard and Statistics stimulate me to teach more efficiently.**
I2: **The Course Dashboard and Statistics stimulate me to teach more effectively.**

Q	strongly disagree 1	2	3	4	5	6	7	8	9	strongly agree 10
D1	0 / 0.00%	1 / 4.35%	2 / 8.70%	2 / 8.70%	1 / 4.35%	2 / 8.70%	0 / 0.00%	5 / 21.74%	5 / 21.74%	5 / 21.74%
D2	1 / 4.35%	1 / 4.35%	1 / 4.35%	2 / 8.70%	3 / 13.04%	0 / 0.00%	3 / 13.04%	2 / 8.70%	4 / 17.39%	6 / 26.09%
AR1	0 / 0.00%	1 / 4.35%	1 / 4.35%	0 / 0.00%	2 / 8.70%	2 / 8.70%	3 / 13.04%	6 / 26.09%	4 / 17.39%	4 / 17.39%
AR2	0 / 0.00%	1 / 4.35%	1 / 4.35%	5 / 21.74%	2 / 8.70%	3 / 13.04%	5 / 21.74%	4 / 17.39%	1 / 4.35%	1 / 4.35%
AR3	0 / 0.00%	1 / 4.35%	2 / 8.70%	1 / 4.35%	2 / 8.70%	2 / 8.70%	4 / 17.39%	8 / 34.78%	2 / 8.70%	1 / 4.35%
AR4	1 / 4.35%	0 / 0.00%	3 / 13.04%	0 / 0.00%	2 / 8.70%	5 / 21.74%	2 / 8.70%	7 / 30.43%	2 / 8.70%	1 / 4.35%
I1	1 / 4.35%	1 / 4.35%	1 / 4.35%	1 / 4.35%	1 / 4.35%	7 / 30.43%	3 / 13.04%	5 / 21.74%	1 / 4.35%	2 / 8.70%
I2	0 / 0.00%	1 / 4.35%	2 / 8.70%	0 / 0.00%	1 / 4.35%	5 / 21.74%	4 / 17.39%	5 / 21.74%	3 / 13.04%	2 / 8.70%