

An Integrated Learning Analytics Approach for Virtual Vocational Training Centers

Peter de Lange*, Alexander Tobias Neumann, Petru Nicolaescu, Ralf Klamma

RWTH Aachen University, Aachen (Germany)

Received 3 November 2017 | Accepted 22 January 2018 | Published 15 February 2018



ABSTRACT

Virtual training centers are hosted solutions for the implementation of training courses in the form of e.g. Webinars. Many existing centers neglect the informal and social dimension of vocational training as well as the legitimate business interests of training providers and companies sending their employees. In this paper, we present the virtual training center platform V3C that blends formal, certified virtual training courses with self-regulated and social learning in synchronous and asynchronous learning phases. We have developed an integrated learning analytics approach to collect, store, analyze and visualize data for different purposes like certification, interventions and gradual improvement of the platform. The results given here demonstrate the ability of the platform to deliver data for key performance indicators like learning outcomes and drop-out rates as well as the interplay between synchronous and asynchronous learning phases on a very large scale. Since the platform implementation is open source, results can be easily transferred and exploited in many contexts.

KEYWORDS

Virtual Vocational Training, Learning Analytics, Digital Blended Learning.

DOI: 10.9781/ijimai.2018.02.006

I. INTRODUCTION

THE digital transformation is affecting vocational training as any other business these days. A virtual vocational training center is a hosted solution for the implementation of training courses in the form of e.g. Webinars. There are many undeniable advantages coming with the utilization of virtual training centers, both for training providers and trainees. Training providers can save the rental for training facilities including the procurement of furniture, media technology and training materials. Furthermore, they can expand their business to other regions within their country or even across borders, e.g. in the European Union. Scaling their business is important for many training providers, in case their business is threatened by economic up- and downturns or by changing demands in job descriptions. They can enter other markets, e.g. translate electronic training materials, adjust materials to different job descriptions or can cooperate with other training providers in larger virtual training centers at low front-off costs. Trainees can save travel time and costs and take parts in training programs offered outside their living areas. However, the virtualization of training has also downsides, most important the loss of social interaction with other learners and the informal exchange of information before, during and after the training. So, the goal of a virtual training center is not only the formal implementation of training courses with the necessary training, testing and certification procedures but also a re-establishment of informal and social learning opportunities.

In this paper, we present the “Virtus Virtual VET Center” (V3C) concept and its implementation on a Web-based platform. It offers

Webinar style training courses with the facilitation of trainers in a synchronous manner as well as self-regulated, asynchronous learning spaces, where learners can learn in a self-paced, socially-aware way together with other learners or alone. To evaluate the impact of the different learning modes on the learning outcomes we developed and applied several learning analytics methods, utilized and combined in a way to get insights into both synchronous and asynchronous learning modes. The platform itself uses different data collection methods based on the different interaction possibilities and also different data analysis and visualization procedures, combining them to an integrated holistic learning analytics approach for vocational training centers. The data and its visualization are accessible both for trainers and trainees, so that they can reflect on their learning progress.

The long-term goal of our research is to blend synchronous Webinar or MOOC style learning with informal social learning in vocational training in a flexible Web-based platform together with the fulfillment of business goals of vocational training centers as well as learning goals of small and medium sized companies. This blending is supported by using our integrated learning analytics approach, which makes it possible to track the learners’ progress throughout the different learning phases and enables both learners and tutors to see differences and similarities in the learning outcomes.

The remainder of this contribution is structured as follows. Section 2 describes the concept and realization of the V3C platform. We then start Section 3 with stating our research questions, describing our data sources and the method of analysis we followed in our evaluation. The section is continued by presenting our results and their interpretation. Finally, Section 4 describes work related to this contribution and Section 5 concludes our paper and provides an outlook on future work.

* Corresponding author.

E-mail address: lange@dbis.rwth-aachen.de

II. A STATE-OF-THE-ART VIRTUAL VOCATIONAL TRAINING CENTER

A state-of-the-art virtual vocational training center has to combine a Learning Management Systems (LMS) with a Personal Learning Environments (PLE) [1]. The LMS has to be designed to be usable by all stakeholders responsible for course (content) creation, thus it has to have a very low technical entry barrier. The PLE should allow for both synchronous, class-style learning and asynchronous, self-regulated learning at the same time to cope with the requirements that vocational training brings. While classical teaching via Webinars is needed in this domain, vocational training needs to take additionally into account asynchronous learning. Since many students are working full time at different companies, they are connected to each other more in the form of a Community of Practice (CoP) [2]. Often, it is impossible for them to participate in each teaching session. To develop a platform that supports both learning modes, the platform needs to have a live video chat functionality as well as static content that is available all the time, like slidesets, videos and online assessment tools that provide feedback to both learners and tutors about the students' progress. Additionally, the system should track the learning progress of the students in form of data collections in the background. This data and its aggregations, clustering and evaluation allows for Learning Analytics (LA [3]) that can provide valuable insights for both students and tutors to either improve their learning or teaching (material) accordingly.

In the following, we present the realization of the "Virtus Virtual VET Center" (V3C), developed in the scope of the European Erasmus+ project VIRTUS. We first presented this platform in [4] and here we only want to give the reader a short overview of the platform's capabilities. We realized our platform as a hyper learning environment consisting of a LMS and a PLE.

Regarding the technical aspects, we used well-established Web development languages and protocols, such as PHP, JavaScript, Java RESTful microservices, HTML5, XMPP and WebRTC. The LMS allows for a modeling-like drag-and-drop design of course rooms (see Fig. 1). This eases the creation of courses, especially for non-technical learning designers, since it shows the learning room already in a "What-You-See-Is-What-You-Get" (WYSIWYG) fashion. Each course is divided first into several modules, which again are divided into multiple learning units.

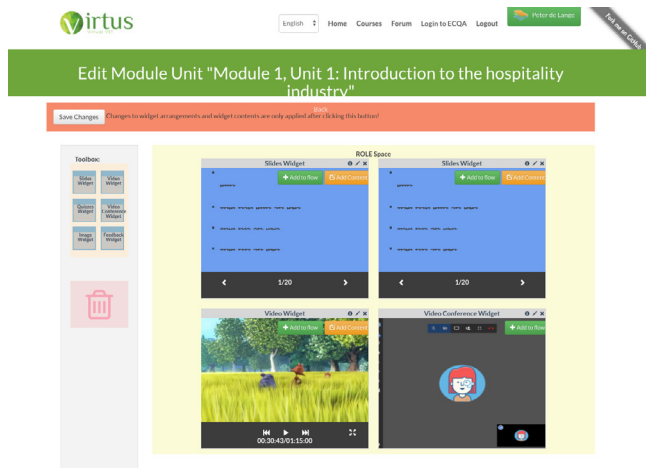


Fig. 1. The modeling view of the V3C platform.

The courses generated by the LMS of our platform are represented as "learning spaces", realized using the "Responsive Open Learning Environment" (ROLE) platform [5]. Fig. 2 shows such a learning

space of a module of the course "Social Entrepreneurship". Each space represents a designed module, with its individual learning units being represented as "learning activities" of the space. This realizes a separation between the individual units which are part of a course, enables activity and progress tracking for individual units and allows for assessment via quizzes of the respective units' learning outcomes. V3C users can autonomously join spaces via the respective course unit in their LMS. Each course unit may consist of video chat, slide presentations, various multimedia content such as audio recordings, videos and images, and self-assessment quizzes. All realized software is Web-based and open source.



Fig. 2. A course room of the V3C platform.

Tutors and other learners can intervene in the learning process at any point via video or text chat that is available for each course. Since our target group for both learning designers and learners consists of people from Italy, Austria, Greece and Spain, the V3C is developed with extensive translation functionality, providing opportunities to offer and translate learning units into different languages. Finally, our platform is linked to the European Certification and Qualification Association (ECQA), which then conducts and assesses the final certified exams. For data protection, we use the OpenID Connect standard to feature a unified login for both the LMS and the PLE as well as for the certification process at the ECAQ.

The LMS of the V3C platform has a per-module analytics section where tutors can see usage statistics as well as progress reports of the course participants. While learners use the PLE, usage information is logged into a MariaDB database. We track user interactions as clicks on different HTML elements. We also implemented a routine that measures the time spent on the respective module unit as duration. This routine sends every minute a request that updates the time. All visualized information shown on the platform are real-time analytics. The analytics section is split up into three different views (participants, feedback and activity). The first view contains information about individual participant results in the learning module selected. It shows the total number of participants of the module and lists them. Since during the creation progress of a module and its units, the tutor has specified the final test of the course, we weight these points with a factor, resulting in a minimum duration the user has to interact with the platform to have a "full completion rate" of the course. Another information to be observed in this view is the assessment of the module. At the end of each module, the user has to complete a quiz. By clicking on a participant, the tutor can inspect the monitored data, split up into the module's units. It now also shows the previously described data of each unit, as well as how much of the quiz was completed and how many correct answers were given. In the second view, tutors can see feedback of the learners. The feedback widget is optional for the learners and serves as a means for the learners, especially when using the platform in an asynchronous, self-regulated learning mode, e.g. to get in contact with the course designer to mention problems with the content. Of course, it can also be used

to submit answers to given tasks by the tutor. The evaluation then has to be done manually by the tutors. The last view of the analytics section displays the “activity graph” of the module. All interactions tracked by the system and the overall activity of the module is displayed here. These visual analytics support enables tutors to track the activity of their modules over time, being able to see at what times students engage themselves most in the platform, which courses are more frequently visited etc. The graph is auto-scaled and shows the whole activity since the beginning of the module, but it also offers the option to select a desired time period. We show a screenshot of these graphs later in the evaluation section in Fig. 3.

III. MONITORING A PERSONAL LEARNING ENVIRONMENT

In the following we describe the evaluation of the usage of the V3C platform. Part of what we present here, the questionnaire results, is based on the technical report created by the VIRTUS project consortium in [6].

A. Hypotheses

In order to evaluate our collected data and gain insightful conclusion regarding the usefulness of our approach, we formulated two research questions we wanted to answer in our evaluation.

RQ1: How well does platform immanent monitoring of student behavior reflect the learning process and how well does the data collection work?

This question for one aims at evaluating the effectiveness of our implemented solution regarding its technical capabilities of dealing with learner-generated amounts of big data. Second, it aims at identifying its capability to successfully make predictions regarding future assessment results, based on the learners’ activity, with a special focus on at-risk students.

RQ2: How does asynchronous, self-regulated learning perform, compared to synchronous, Webinar-driven learning modes, if applied in the same platform?

Here, we want to compare the two contrary learning modes to evaluate their differences in final assessment results, drop-out rates and activity / engagement with the platform. Since our clustering of students into synchronous and asynchronous “evaluation groups” is done at random, our question is not aiming at categorizing students into different “types of learners”, but at evaluating, what effect on the learning outcomes can be seen when replacing mandatory, tutor-supported synchronous learning sessions with additional time for self-regulated learning. Since our platform allows the evaluation of the two groups in the same learning rooms, we can achieve comparable results.

B. Data Sources

The data used in this evaluation was gathered from the platform immanent learning analytics mechanism, additionally complemented with the use of a questionnaire that addressed the general impressions participants had while using the platform. Our evaluation participants were either already workers in the tourism sector or planned to enter this sector. In fact, the courses offered by the platform concerned the two domains of (*Social Entrepreneurship* and *Tourism and Hospitality*). We had a total of 114 learners from four different countries, namely Austria, Italy, Greece and Spain. The evaluation period span two months. The synchronous learning phase was performed in the *Tourism and Hospitality* course and covered 18 units in four modules. Every module consisted of at least two units, with a maximum of up to five units. We were able to recruit 15 learners for the synchronous and 72 for the asynchronous evaluation.

Each of the four vocational training Webinars spanned two hours and covered one module of the course, resulting in a four-day streak of consecutive synchronous evaluation sessions. The tutor facilitating the synchronous learning session spoke English and used the platform-immanent video conference widget. Starting with the beginning of the first synchronous evaluation session, learners who participated in this phase had two weeks to complete the course. The asynchronous learning phase began simultaneously with the synchronous learning phase, but here learners had 60 days to complete the course. Table I shows the number of learners who participated in each module. Finally, we invited the participants to fill out a questionnaire. We received 48 submissions here, without clustering them into the two evaluation clusters.

TABLE I. PARTICIPANTS PER MODULE

Module	Synchronous Participants	Asynchronous Participants
SE1	0	42
SE2	4	50
SE3	0	37
SE4	0	37
SE5	3	34
TH1	15	72
TH2	14	58
TH3	15	55
TH4	14	54
TH5	9	51

(SE = Social Entrepreneurship, TH = Tourism & Hospitality)

C. Methods of Analysis

For collecting and public provision of our learning analytics, we used the MobSOS monitoring concept of our research group’s Community Information System (CIS) platform [7]. The MobSOS Query Visualizer (MobSOS QV) was used to embed multiple measures together in a dashboard-like fashion, which we provide publicly on the platform. The stored queries for those visualizations are parameterized and can be used for every module and unit. For our analysis, we used the user activity and the scored result of the user. Currently, an activity is defined as any click interaction of the learner with the PLE, e.g. click on any button, watch a video or answering a quiz.

During the synchronous evaluation sessions, participants were allowed to use the whole platform. Although these sessions were conducted in English, participants also at this time were able to use the platform in their desired language. After each session, the tutor manually created an attendance list, which marked the attendants as part of the synchronous evaluation group, allowing us to cluster the participants into an asynchronous and a synchronous cluster for evaluation.

Both asynchronous and synchronous learners used the same learning rooms simultaneously. A module was marked as completed when the participant completed the quiz at the end of it. We only considered learners who have participated in each module of the course for the comparison of the scored result and the activity in the PLE. The scored result of a learner is the percentage of the correct questions of all modules. The activity is the percentage of all activities done by a user in the whole course compared to the activities done by the learner who has done the most activities in this course. We base the calculation of the drop-out rate by only taking those participants who finished all previous modules of a course, although our platform offers the possibility to only attend single modules of a course without the need to attend a previous module.

For further data which could not be derived by platform immanent monitoring, we follow the “observe, where possible; only survey, where inevitable” approach [7] by using additional five-level Likert

items for yet missing, however relevant subjective factors. The questions of the survey were translated into the participants' native languages and were handed out in digital or printed format to the participants of both phases. Questions regarded both the platform as well as the teaching material, and we focused on those that deal with user satisfaction of the system since this is part of the success awareness. We had no means to link the questionnaire to individual persons, nor did we cluster them into asynchronous and synchronous participants' answers.

While using the platform, all participants were aware of the monitoring of their progress and behavior by our system. They had the possibility to always review their personal data we collected, such as quiz results and the time they spent in a learning space. Only the overview of all participants of a module was restricted to the tutor / learning designer as well as the visualization of the aggregated and anonymized learning room activity. This "Module Activity" view is also depicted in Fig. 3 of the next section.

D. Results

In this section, we present the results of our evaluation. It has to be noted that we only consider the course "Tourism and Hospitality" for analysis since it is the only course where we have both a synchronous and asynchronous evaluation, which enables the comparison between the two learning modes.

As already mentioned in the previous section, our platform features a visual analytics feature for tutors, teachers and learning designers to monitor the activity of their modules over time. Fig. 3 shows this view for all four modules of the course. It displays the absolute number of activities for each module over time. The timespan in this figure was chosen to match the sixty-day evaluation period. It can be seen that two peaks in every module exist, which correspond to the final assessment of both the synchronous evaluation phase (after 14 days) and the one of the asynchronous evaluation around six weeks later. As it can also be observed, there is a smaller, yet considerably high activity over the whole evaluation cycle, which is an indicator for continuous usage of the platform by the participants during our evaluation.

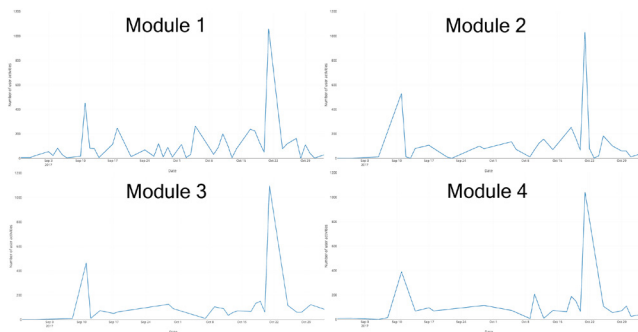


Fig. 3. Module activity over time.

Our next analytics compared the participants activity with their scored result in the assessment. Fig. 4 shows this comparison. Here, we used relative values for both activity and scored result to make them comparable, due to the different number of participants in both evaluation phases. We only considered those students that took part in the whole course. The Pearson correlation coefficients over all four modules were $p = 0.4593$ for the asynchronous phase and $p = 0.3589$ for the synchronous phase.

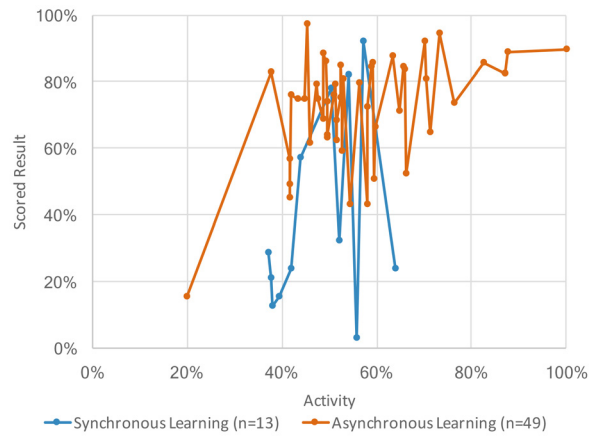


Fig. 4. Activities compared to scored result.

Fig. 5 gives an overview about the final aggregated results of the participants for both evaluations. Again, we use relative values here to make the results comparable and removed all prior drop-out participants from this statistic.

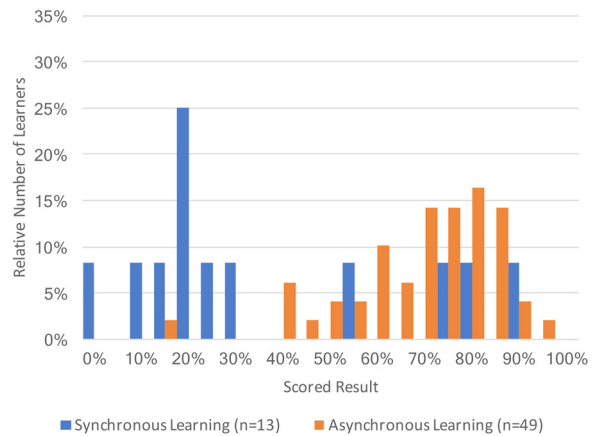


Fig. 5. Assessment results.

Our last analytics concerned the drop-out rate of both synchronous and asynchronous evaluation participants. Fig. 6 shows this statistic. The percentages are always relative to the base number of participants ($n = 15$ and $n = 72$). As it can be seen, here we have the complete participants of both evaluation groups of the evaluated course. When counting together the drop-outs of this figure, the number shrinks to the number of the prior three analytic measures ($n = 13$ and $n = 49$).

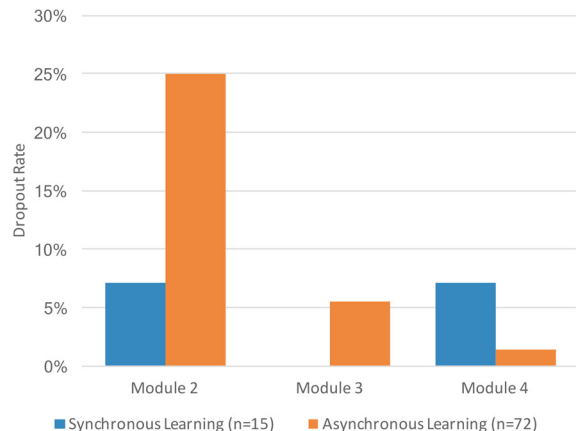


Fig. 6. Drop-out rate.

Finally, Fig. 7 shows the result of the questionnaire. We aggregated both synchronous and asynchronous participants together and also took into account learners from the other course (“Social Entrepreneurship”) which we only used for asynchronous evaluation and thus have not taken into account for the prior analytics presented in this section.

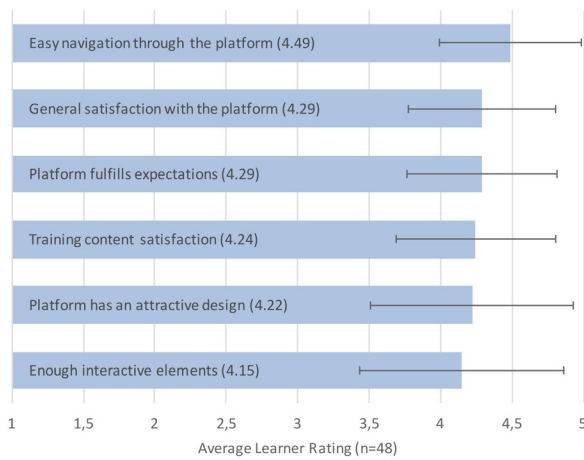


Fig. 7. Results of questionnaire.

E. Interpretation

With our evaluation, we wanted to answer two research questions. RQ1 regarded the effectiveness of our technical monitoring solution as well as its capability to make predictions regarding the learning outcome. For the first part, we can state that we were able to monitor all data that was produced by the learners in the evaluation and our platform was capable of handling both synchronous data streams of multiple people collaborative using the same learning room, as well as that it was robust against a long-term evaluation of two month for the asynchronous evaluation phase. This is also backed by Fig. 3, which shows a constant activity flow with the peaks clearly representing assessment deadline phases. For the second part of RQ1, we have to be a little more cautious with our interpretation. Fig. 4 shows a rather clear correlation between activity in a learning room and scored assessment results for participants of the asynchronous learning cluster. The Pearson correlation also confirms this with its value close to 0.5, which is considered as rather “correlation confirming” by most literature. For the asynchronous learning phase though, the relation is not that clear. The Pearson correlation again shows there is a slight correlation between both values, but the graph does not really allow for a clear prediction capability. Reasons for this might be the smaller subset of the asynchronous cluster. With a larger amount of participants here, the results might have been clearer. We also tried to evaluate the results on a per-module basis (instead of the four-module aggregated view), but these present similar results.

For RQ2, which asked how synchronous and asynchronous learning modes perform if applied in the same platform, first we can say that there is a clear trade-off between time provided to the participants and resulting assessment results. The cluster of asynchronous learners, which had about 75% more time to complete the assessments, scored about 33.38% better results than the cluster of synchronous learners. This can also be seen in Fig. 5. On the other hand, it has to be said that reducing the time needed to train people by two-third results also in a very high reduction of costs needed to be spend by employers to provide the training to their employees. Ultimately, it has to be decided by the facilitators of vocational training courses, if this trade-off between the time needed to train people and the expected results is worth of taking.

Regarding user acceptance of the platform, our questionnaire results (see also Fig. 7) indicate a high satisfaction of the users with the platform itself. Since all questions received similar high ratings, we will not discuss them in detail, but it has to be said that these ratings might also result from the domain we chose to evaluate our platform. People working or planning to work in the domain of tourism, especially on a service level, might not be used to the support their training by the means of multi-media in general. So, the availability of a platform that provides them with the possibility to perform their training courses online might have been perceived as very appealing, resulting in these high ratings of user acceptance in terms of learners.

F. Limiting Factor

Performing synchronous evaluations that span multiple days is a resource consuming task. Especially, when the participants have to be recruited from a particular vocational domain, with full-time job schedules and limited time to spend in training that is not directly financed by their employer. Therefore, we limited the comparison between the two learning modes presented in the previous section to the “Tourism & Hospitality” course, and covered only four out of five modules here. We decided to leave out the duration measure, which represents the time spent for a module, due to some technical uncertainties we had with the results they produced. This is definitely something we need to rethink and reimplemented for further evaluations. Currently, our assessment is limited to the usage of multiple choice quizzes. We are aware of the shortcomings of this approach regarding the capabilities of multiple choice questions to assess learning success as a whole. For the future, we are working on integrating the results of the ECVET certified final exams of the courses, done by the ECQA, into our dataset. For the moment, we are not able to consider those here due to some restrictions. During the analysis, we noticed that two learners of the asynchronous phase participated twice in a course, thus we removed both datasets of the participants from the evaluation data. One thing we did find out when perform our analytics was the following. It would have been interesting also to have a cluster of learners that had the same time to finish the course as the synchronous learners had (two weeks), but without providing them the opportunity to participate in the Webinar. Finally, we want to state that our user acceptance questionnaire only aimed at evaluating the platform from a learner’s perspective. This results from the fact, that the platform was developed as part of the same project (VIRTUS), where also the VET trainers that provided the learning content via the platform were partners of. Thus, we had no external VET providers that we could have interviewed to find out about their opinions regarding the platform, although we deem this an area worth of further research.

IV. RELATED WORK

Vocational training adapted to the digital, virtual domain not as fast as e-learning conducted in higher education at university level for example. Although its potential was first mentioned already at the end of the last century [8], its adoption is still in its infancy. The result of this is that studies performing learning analytics, to the best of our knowledge, consider either university/college or schools as the educational setting [9]. From a technology point of view, most learning analytics approaches use data generated by Massive Open Online Courses (MOOCs) created by open source course management system, with the most famous one being “Moodle” [10]. Here, using standardized data formats [11] to apply learning analytics on, like the Learning Record Store (LRS) of the xAPI standard [12], is an active research topic. Having a well-defined data (exchange) format allows for an easier handling of analyzing data, but it does not solve the question what data needs to be collected to understand past and predict future student behavior and react accordingly. Here, techniques

like Educational Process Mining (EPM) can be used to retrospectively derive meaning of collected data logs. Bogarín et al. [13] have used this technique together with clustering to predict success rates of students taking a course of a Moodle course management system installation. Click-based analysis is also very often applied to learning analytics in virtual learning environments, like for example the authors of [14] used the GUHA (General Unary Hypothesis Automaton) method, a data mining technique, to predict student drop-out rates. In general, we can conclude that most studies consider the prediction and improvement of assessment results, activities and drop-out rates [15] as their main goal. While the first two are of equal importance in asynchronous vocational training, the drop-out rate problem is more bound to the MOOC context, since most vocational training still is done in a teaching-style manner, with often mandatory participation. Without self-regulated learning, drop-out rates in this setting are similar to classroom-like situation, which also corresponds to our study results. This especially lays the focus on the study of self-regulated learning with learning analytics to identify and analyze self-regulated learning strategies and how to improve them [16].

It appears that there is still a lack of research for virtual vocational training, especially in the domain of learning analytics. To the best of our knowledge, there exists no research that compares the effects of asynchronous and synchronous learning virtual vocation training and education.

V. CONCLUSION & OUTLOOK

We presented an integrated learning analytics approach that allows to compare asynchronous and synchronous learning phases and draw conclusions for certification, interventions and gradual course improvement. It is embedded into a technical platform for the creation, implementation, deployment and performance of virtual training courses. We evaluated our contribution with two vocational training courses, one for social entrepreneurship and one for tourism & hospitality. Based on an English version, we translated both courses in four languages (German, Italian, Greek and Spanish) and conducted them with 114 learners from different European countries. Our main focus for this paper was the interplay of synchronous and asynchronous learning phases to demonstrate the ability of the platform to blend Webinar-style course units with phases of self-regulated learning on the platform. Additionally, we had an eye on important questions like for both virtual training centers offering courses as well as small and medium sized companies sending their employees for vocational training, e.g. drop-out rates.

The interpretation of our learning analytics results shows that the combination of Webinars and self-regulated learning saved additional resources beyond the pure availability of a virtual training center. This is the first evaluation of the learning platform and the evaluation of asynchronous learning phases is still ongoing at the time of the writing. Moreover, we have not differentiated yet the performance of learners in different language versions of the same course. That means that the possibilities for advanced learning analytics are not yet explored and we can expect more and better feedback from the platform and its analytic capabilities. What we can say for now is that the platform scales well both in the number of courses and learners, since the underlying ROLE framework has been in use for self-regulated learning since many years and one installation is capable of serving several thousands of learners. In further research, we will focus also on community learning analytics for understanding the social processes in virtual training centers better from an empirical perspective.

ACKNOWLEDGMENT

This research was funded by the European Commission under the European Union's Erasmus+ project "VIRTUS" (grant no. 562222).

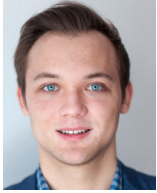
REFERENCES

- [1] M. van Harmelen, "Personal Learning Environments", in Proceedings of the 6th IEEE International Conference on Advanced Learning Technologies, Washington, DC, USA: IEEE Computer Society, 2006, pp. 815–816.
- [2] E. Wenger, *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, UK: Cambridge University Press, 1998.
- [3] R. Ferguson, "Learning analytics: drivers, developments and challenges", in *International Journal of Technology Enhanced Learning*, vol. 4, no. 5-6, pp. 304–317, 2012.
- [4] P. de Lange, P. Nicolaescu, and R. Klamma, "VIRTUS Virtual VET Centre (V3C): A Learning Platform for Virtual Vocational Education and Training", in *European Conference on Technology Enhanced Learning*, 2017, pp. 500–503.
- [5] D. Renzel, R. Klamma, M. Kravčík, and A. Nussbaumer, "Tracing Self-Regulated Learning in Responsive Open Learning Environments", in *Advances in Web-Based Learning - ICWL 2015: 14th International Conference*, Guangzhou, China, November 5-8, 2015, Proceedings, Berlin-Heidelberg: Springer, 2015, pp. 155–164.
- [6] RWTH Aachen University, "R2.3: State of the Art Report on Virtual Organizations in Education and Training: Virtus Project Technical Report", 2016.
- [7] D. Renzel, R. Klamma, and M. Jarke, "IS Success Awareness in Community-Oriented Design Science Research", in *New Horizons in Design Science: Broadening the Research Agenda: 10th International Conference, DESRIST 2015*, Dublin, Ireland, May 20-22, 2015, Proceedings, Switzerland: Springer International Publishing, 2015, pp. 413–420.
- [8] R. C. Shank, *Virtual learning: A revolutionary approach to building a highly skilled workforce*. New York: ERIC; McGraw-Hill, 1997.
- [9] N. Nistor, M. Derntl, and R. Klamma, "Learning Analytics: Trends and Issues of the Empirical Research of the Years 2011-2014", in *Design for Teaching and Learning in a Networked World: 10th European Conference on Technology Enhanced Learning, EC-TEL 2015*, Cham: Springer International Publishing, 2015, pp. 453–459.
- [10] M. Dougiamas and P. Taylor, "Moodle: Using learning communities to create an open source course management system", in *World Conference on Educational Multimedia, Hypermedia and Telecommunications: EDMEDIA'03*, 2003.
- [11] Á. Del Blanco, Á. Serrano, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, "E-Learning standards and learning analytics. Can data collection be improved by using standard data models?", in *Global Engineering Education Conference (EDUCON)*, 2013 IEEE, 2013, pp. 1255–1261.
- [12] J. M. Kevan and P. R. Ryan, "Experience API: Flexible, decentralized and activity-centric data collection", *Technology, Knowledge and Learning*, vol. 21, no. 1, pp. 143–149, 2016.
- [13] A. Bogarín, C. Romero, R. Cerezo, and M. Sánchez-Santillán, "Clustering for improving educational process mining", in *Proceedings of Learning Analytics and Knowledge Conference 2014, LAK '14*, New York: ACM, 2014, pp. 11–15.
- [14] A. Wolff, Z. Zdrahal, A. Nikolov, and M. Pantucek, "Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment", in *Proceedings of the Third International Conference on Learning Analytics and Knowledge*: ACM, 2013, pp. 145–149.
- [15] J. L. Santos, J. Klerkx, E. Duval, D. Gago, and L. Rodriguez, "Success, activity and drop-outs in MOOCs an exploratory study on the UNED COMA courses", in *Proceedings of Learning Analytics and Knowledge Conference 2014, LAK '14*, New York: ACM, 2014, pp. 98–102.
- [16] J. Maldonado-Mahauad, M. Pérez-Sanagustín, R. F. Kizilcec, N. Morales, and J. Munoz-Gama, "Mining theory-based patterns from Big data: Identifying self-regulated learning strategies in Massive Open Online Courses", *Computers in Human Behavior*, vol. 8, pp.179-196, 2017.



Peter de Lange

Peter de Lange received his Master's degree in computer science in 2015 from RWTH Aachen University. Peter works as a research assistant and PhD student within the Advanced Community Information Systems group at the Information Systems and Databases chair, RWTH Aachen University, Germany since 2015. He is involved in several EU projects in Technology Enhanced Learning as well as research and teaching. Among his research interests are Web engineering, community information systems, learning analytics and peer to peer infrastructures.



Alexander Tobias Neumann

Alexander Tobias Neumann received his Bachelor's degree in computer science from the RWTH Aachen University in 2016. In his thesis he focused on human-computer interaction with emphasis on tangible user interfaces. Currently, he is working towards his Master's degree and focuses on technology-enhanced learning.



Petru Nicolaescu

Petru Nicolaescu works as a PhD student within the Advanced Community Information Systems group at the Information Systems and Databases chair, RWTH Aachen University, Germany. He is involved in major impact EU projects in technology enhanced learning, research and teaching since 2012. He received his engineering (B.Sc.) degree in computer science in 2009 from Politehnica University Bucharest and his Master degree in Software Systems Engineering in 2012 from RWTH Aachen University. Among his research interests are information systems and infrastructures, collaboration technologies, Web engineering and social network analysis.



Ralf Klamma

Ralf Klamma holds diploma, doctoral and habilitation degrees in computer science from RWTH Aachen University. He leads the research group "Advanced Community Information Systems" (ACIS) at the information systems chair, RWTH Aachen University. He is known for his work in major EU projects for Technology Enhanced Learning (PROLEARN, GALA, ROLE, Learning Layers, TELMAP, Tellnet, CUELC, SAGE, BOOST, VIRTUS and WEKIT). Ralf organized doctoral summer schools & conferences in Technology Enhanced Learning and Social Network Analysis. He published more than 150 scientific papers. He is on the editorial board of Social Network Analysis and Mining (SNAM) and IxD&A. His research interests are community information systems, serious games, augmented reality, web engineering, social network analysis, requirements engineering and technology enhanced learning.