DERIVING LEARNER-CENTRIC PLATFORM FEATURES THROUGH CUSTOMER REVIEW MINING

Completed Research Paper

Timo Grüneke, FIM Research Center, University of Bayreuth, Branch Business & Information Systems Engineering of the Fraunhofer FIT, Bayreuth, Germany, timo.grueneke@fim-rc.de

Abstract

In an age where digital platforms are revolutionizing education, our study critically examines MOOCs, focusing on the nuanced interplay between platform product features and learner satisfaction. By analyzing 142,976 Coursera reviews using natural language processing (NLP) and supplementing this with data from various platforms, we uncover specific platform features to research learner satisfaction in the future. This deep dive, utilizing Latent Dirichlet Allocations (LDA) for topic modeling, reveals insights into what learners' value in their online education journey. Thereby, we identify 32 essential features, categorized into seven structural characteristics. This research not only fills a significant gap in understanding the crucial role of platform features in shaping learner satisfaction in MOOCs but also offers insights for platform providers to act on. These insights are key to enhancing the digital learning experience, ensuring that platforms meet and exceed the evolving expectations of learners in this dynamic educational landscape.

Keywords: MOOC, Topic Modeling, Learner Satisfaction, Digital Education.

1 Introduction

In the rapidly evolving digital age, platforms are powerful tools for connecting users, providing services, and promoting social and economic interactions (Hein et al., 2020; Sun and Gregor, 2023). Likewise, digital platforms have reinforced the transformation in education and enabled the advent of massive open online courses (MOOCs) (Hew and Cheung, 2014). Platforms like Coursera, edX, and Udacity democratize access to knowledge as digital learning supported by information systems opens new opportunities to anyone with an internet connection to access global knowledge, ideas, expertise, and qualifications (Blayone et al., 2017; Hone and El Said, 2016). Further, using platforms as tools for technology-mediated learning (Alavi and Leidner, 2001) the underlying technological infrastructure allows scaling the learning experience for self-paced (Hone and El Said, 2016; Wambsganß et al., 2021) and self-regulated learning (Dever et al., 2022). Following this trend, the number of courses on learning platforms increased heavily, resulting in over 220 million students and over 150 thousand courses offered worldwide (excluding China) (Class Central, 2023).

Today, learning platforms are not mere repositories to host and display content (Islam, 2012) but two-sided markets or ecosystems where instructors offer online learning opportunities, and learners are enabled to engage with their peers, acquire knowledge, and seek lifelong learning (Islind et al., 2021; Williamson, 2021). Despite the vast content and numerous courses available, not all learners find their experiences on these platforms fulfilling. The crux lies in identifying what features significantly contribute to or detract from learner satisfaction. Understanding these critical features is essential as satisfaction determines the user's continuance intention (Bhattacherjee, 2001). Thus, the platform must acquire an understanding of the learner to satisfy their customer needs (Hew et al., 2020). Beyond that, the platform

must identify and adopt features that delight customers to yield competitive advantages (Matzler et al., 1996) recognizing learning platforms as complex sociotechnical systems (Williamson, 2021).

Precisely because education is a pillar in individual and societal development understanding the nuances of satisfaction is of paramount importance. We see a surge in research surrounding online education, where most studies emphasize pedagogical strategies or instructional design (Sweller, 2020), data analysis (Romero and Ventura, 2020), or technological aspects (Zawacki-Richter et al., 2019). However, the intersection of platform features and learner satisfaction in MOOCs has been inadequately explored (Jung et al., 2019), with studies typically isolating individual items like targeted teaching of competencies through suitable programs (Abidin, 2021) or the specific attributes for designing a learning video (Mayer et al., 2020). Particularly in the context of MOOCs, only a few studies truly dive into platform features and learner satisfaction (Hew et al., 2020) and grasp the underlying market mechanisms of platforms (Williamson, 2021). Thus, there is a conspicuous need for a study that synthesizes these disparate features into a cohesive framework. Hence, we aim to identify platform features from learners' course reviews that platforms can use to improve learner satisfaction. Furthermore, our goal is to construct a methodological approach leveraging Natural Language Processing (NLP) and qualitative evaluation methods to unobtrusively uncover actionable insights. We hereby answer the call of Li et al. (2020) to use NLP to create domain-specific knowledge and also follow the efforts of Reuver et al. (2018) to examine digital platforms across different industrial contexts. Additionally, we seek to contribute to the discussion on the advancing maturity of data mining in education (Romero and Ventura, 2020). Against this backdrop, we ask:

RQ: *Which platform features can learning platform providers derive from the voice of learners expressed in course reviews to ensure customer satisfaction?*

To answer our research question, we first use NLP to understand learners' different perspectives on MOOCs (Berger et al., 2020). We collect and analyze 142,976 Coursera reviews using topic modeling to monitor the learner's opinions (Peng and Xu, 2020) and to identify platform features (Hew et al., 2020; Rice, 2006). Employing Latent Dirichlet Allocations (LDA) (Blei et al., 2003), we uncover topics that learners talk about regarding MOOCs. Second, we additionally leverage purposefully sampled data from organizational documents, blog posts, and textual website data from four additional learning platforms (Chun Tie et al., 2019). Third, we adapt the lens of the Kano method (Kano et al., 1984) to guide the process of feature identification. The method classifies and prioritizes customer needs based on their impact on satisfaction focusing on product or service features (Chen et al., 2022; Gimpel et al., 2018). Since product features are an important prerequisite for the Kano model (Griffin and Hauser, 1993), we thoroughly evaluate the product features (Sonnenberg and vom Brocke, 2012). As a result, we identify 32 features clustered into seven structural characteristics. These findings aim to assist platform providers in aligning their offerings with user preferences providing a foundation for future research in this area.

Our paper is organized as follows: Initially, we explore technology-mediated learning, MOOCs, and learner satisfaction. We then detail our research design for gathering and analyzing data. Next, we present our findings, highlighting 32 identified features and their characteristics. Following this, we compare our results with existing studies. The paper ends by outlining theoretical contributions to guide further research and practical recommendations for platform providers.

2 Theoretical Background

2.1 Technology-mediated learning and Massive Open Online Courses

Technology significantly reshaped the educational landscape (Rabin et al., 2019), creating learning environments needed to learn more effectively (Chang, 2016). This facilitates "an environment in which the learners' interactions with learning materials (e.g., readings, assignments, exercises), peers, and/or instructors are mediated through advanced information technologies" (Alavi and Leidner, 2001, p. 2). With technology as a means to education (Janson et al., 2020) the environment is inherently structured by information systems that guide individuals toward achieving learning outcomes (Gupta and Bostrom,

2009). This shift has been particularly evident with the rise of MOOCs, which offer flexibility, allowing individuals to learn at their own pace, devoid of time and space constraints (Rabin et al., 2019). This development towards technology-mediated learning enhanced the scalability of educational experiences (Williamson 2021), crucial for accommodating a wide array of learning preferences (Hone and El Said, 2016; Wambsganß et al., 2021). Impey and Formanek (2021) observe that the proliferation of courses available on learning platforms has propelled MOOCs into the spotlight, underscoring the pivotal role of technology in expanding educational accessibility. However, understanding MOOC success is a multifaceted construct not yet sufficiently researched (Hew et al., 2020). Simultaneously, high dropout rates and sustaining learner engagement are pressing challenges (Hone and El Said, 2016; Hew and Cheung, 2014). Also the quality remains a contentious issue, suggesting a need for a more nuanced exploration of MOOC platforms (Conache et al., 2016; Martin and Bolliger, 2022). It is precisely for this reason that research focuses on different MOOC aspects. For example, Conache et al. (2016) undertake an analysis of various elements such as the business model, the course experience, and the provision of foreign languages. In contrast, research also focuses on very specific aspects, such as how the effectiveness of instructional videos can be improved (Mayer et al., 2020). Another challenge is the assessment of MOOC success solely through the course completion rate, which is insufficient. Hew et al. (2020) argue for the inclusion of "learner satisfaction" as a more holistic measure recognizing that learners enroll with diverse intentions and goals. Therefore, the design and functionality of MOOCs, including the features that facilitate communication, lectures, study materials, quizzes, assignments, and grading, are critical for enhancing student learning satisfaction (Anggraini et al., 2018; Jung et al., 2019). Thus, the intersection of platform features and learner satisfaction in MOOCs presents an opportunity for further exploration (Jung et al., 2019).

2.2 Satisfaction on technology-mediated learning platforms

Platforms are subject to network effects that positively influence the platform's success (Garcia-Swartz et al., 2019). Hence, learning platforms require active ecosystem governance (Gawer, 2014) to balance different interests (Darking et al., 2008) and fulfill customer expectations. Customer expectations constitute the basis of the Kano model and explains customer satisfaction based on the degree of implementation or availability of certain attributes of products or services (Kano et al. 1984). This helps to pinpoint the characteristics impacting customer satisfaction (Bailom et al., 1996) and in prioritizing software engineering requirements (Mkpojiogu and Hashim, 2016). Ultimately, this allows platforms to gain a competitive advantage through implementing attractive qualities (Berger et al., 1993). However, the characterization of these features as qualities is dynamic, changing over time and following a specific lifecycle (Kano, 2001; Löfgren et al., 2011). This evolving nature of what constitutes quality has significant implications for product and service providers highlighting the need for regular assessments of attributes (Witell et al., 2013). In the educational context, students' satisfaction is reflected through learning satisfaction as a feeling or attitude of learners that their desires and needs can be fulfilled in learning activities or processes (Yu, 2022). Hence, the long-term success of learning platforms, particularly in the context of MOOCs, is heavily influenced by several key factors: the quality of the courses, the learning experience provided by instructors, and the features implemented on the platform (Chowdhury et al., 2022; Papp, 2000; Hew et al., 2020). Additionally, the educational content, the stability of the information system, and the teaching methods employed in online courses play a substantial role in shaping students' motivation to engage with these courses (Chowdhury et al., 2022). Furthermore, satisfaction has emerged as an increasingly recognized success indicator in MOOCs, underscoring the importance of these aspects in contributing to the overall effectiveness and appeal of online learning platforms (Rabin et al., 2019; Hew et al., 2020). Unlike traditional education, MOOCs attract a diverse range of learners with different motivations (Hone and El Said, 2016). Therefore, measuring learner satisfaction offers a more relevant assessment of a MOOC's effectiveness. This focus not only gauges the success of the course but also helps attract more participants, benefiting the platform both financially and reputationally (Clow, 2013; Williamson, 2021). To complement this, various frameworks and models have been proposed, breaking down online learning satisfaction into distinct categories such as

learner, course, instructor, and program and organization (Martin and Bolliger, 2022). However, research on platform features and their impact on learner satisfaction in MOOCs remains underexplored (Hew et al., 2020). An innovative method to gather insights on this aspect is through online reviews, which provide a rich and unobtrusive source of data on user satisfaction (Jin et al., 2019). This approach aligns with the broader trend of utilizing customer feedback to understand preferences and design requirements, shedding light on the 'voice of the customer' in online learning (Yang et al., 2019).

3 Research Design

3.1 Topic modeling and latent dirichlet allocation

The research design of this study is split into two blocks (Figure 1). In the first block, we utilize course reviews to identify eight latent topics learners reported. Our approach profits from reviews reflecting specific information on the learner's experience that would not be accessible otherwise (Büschken and Allenby, 2016). On the other hand, qualitative (e.g., in-depth interviews) and quantitative approaches (e.g., questionnaires) are limited in their capabilities for studying large-scale phenomena (Schmiedel et al., 2019). In this regard, NLP can partly mitigate those limitations by unobtrusively collecting and analyzing naturally occurring data without being prone to biases from self-reported data (Hacker et al., 2020). Thus, we draw on topic modeling using LDA (Blei et al., 2003) guided by Müller et al. (2016), Debortoli et al. (2016), and Schmiedel et al. (2019).

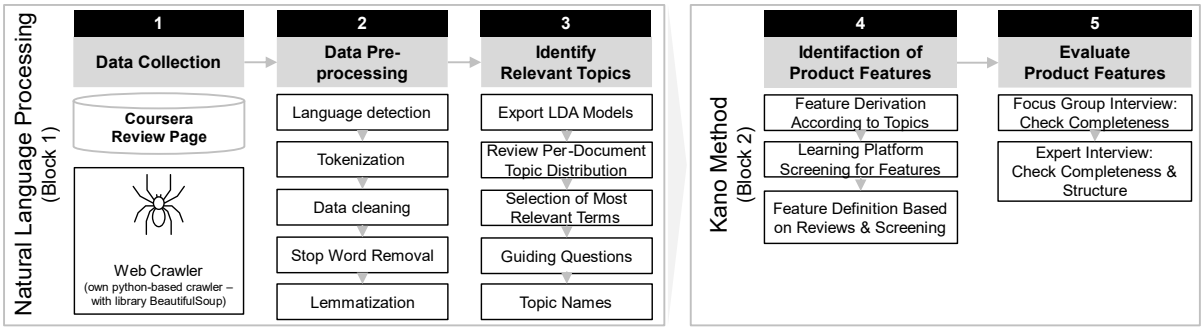

Figure 1. Research Design.

First, we collected review data from Coursera and randomly chose 25 courses from different fields, such as technology, engineering, business, and general management. In March 2022, we used a self-programmed Python-based crawler to collect 142,976 reviews, spanning a timeframe of seven years between August 2015 and February 2022. The number of reviews submitted for each course varied between 2,158 and 9,965 reviews (dataset and the course list can be accessed under the following link: https://doi.org/10.5281/zenodo.10896695). Second, we perform NLP pre-processing steps as textual data lacks well-defined structures and is prone to often include noise (Debortoli et al., 2016). Third, we apply LDA to discover topics in learners' reviews, as the algorithm assigns words and documents to topics based on the principle that "words that occur in the same contexts tend to have similar meanings" (Turney and Pantel, 2010, p. 143). LDA considers texts as mixtures of topics, assigning words to these topics, resulting in topic keywords and the likelihood of reviews relating to them (Blei et al., 2003). To choose the correct number of topics, we employed the semantic coherence metric, aligned with human topic quality judgments (Mimno et al., 2011; Schmiedel et al., 2019), to identify the optimal number, which was eight for our dataset. The results are commonly represented as lists including the top-n most probable words per topic that aid in assessing the significance of a topic (Blei et al., 2003). Building on that, we then systematically examined each topic by analyzing the documents that exhibited a pronounced correlation. This was facilitated through the utilization of per-document topic distribution. Adhering to the guidance of Boyd-Graber et al. (2014), we further employed guiding questions for (qualitative) topic interpretation (Debortoli et al., 2016). This approach aimed to ascertain the meaningfulness,

coherence, and utility of individual topics while also evaluating the appropriateness of the assigned topics to the respective documents.

3.2 Identification of product features

In the second building block, we draw on the Kano method as a theoretical lens of customer satisfaction to identify learning platform features. In step four, we analyzed the eight topics to identify occurring noun phrases resembling product features (Hu and Liu, 2004). Product features are attributes that exist on a continuum of specificness from concrete to abstract (Park et al., 1991). Concrete features are, for example, features of medical health applications like the health record that allows for storing a user's personal and medical information (Gimpel et al., 2021). Therefore, we understand a feature as a specific quality or attribute of a product or service (Kano et al., 1984). We limit the feature set to those features that are already implemented on learning platforms following the basic idea that the identified features enable exploitative innovation to make incremental improvements to existing products and services for the customer (i.e., learners) (Benner and Tushman, 2003; O'Reilly and Tushman, 2008; Božič and Dimovski, 2019) and increase efficiency (Jansen et al., 2008). After compiling the initial feature set, we resort to a purposeful data sampling strategy (Chun Tie et al., 2019) to collect additional data to help answer our research question (Charmaz, 2006). We focus the data sampling on extant texts that are already constructed "because of their relative availability, typically unobtrusive method of data collection, and seeming objectivity" (Charmaz, 2006, p.37). Thus, we searched for organizational documents, blog posts, and textual website data from four additional learning platforms (Chun Tie et al., 2019). This complements the feature set and informs our process of developing feature definitions.

In step five, we evaluate our findings to ensure accuracy, performance, and determine usability (Alturki et al., 2011; March and Smith, 1995). Sonnenberg and vom Brocke (2012) propose various methods and criteria for evaluation. We select interviews as a qualitative research method for evaluation (Creswell and Creswell, 2018; Helfferich, 2014; Myers and Newman, 2007) and select the criteria completeness, comprehensibility, and level of detail to evaluate our feature set (Sonnenberg and vom Brocke, 2012). We split the evaluation into two cycles: the first to evaluate the criteria completeness in focus group interviews and the second to ensure comprehensibility and level of detail in expert interviews.

First, we employed focus group interviews to evaluate our feature set against the criteria completeness. Focus group interviews are used as a qualitative approach to gain a deeper understanding of issues (Merton, 1987; Morgan, 2021). The method does not intend to capture a statistically representative sample of a broader population target but focuses on data collection from a deliberately selected group of individuals (Nyumba et al., 2018). Focus groups' interactive and synchronous group discussion format allows members to debate, agree or disagree with each other's perspectives and expound on previously expressed viewpoints (Nili et al., 2017). Our subjective ending criteria (Nickerson et al., 2013) is to establish that no fundamental features are left out as no additional features for exploitative innovation activities are suggested. We selected a balanced set of participants (active MOOC platform users, area of expertise, and gender) (Table 1), conducted the interviews online (i.e., Microsoft Teams and Mural) and recorded the sessions without rewarding the interviewees for participation. On average, the interviews lasted 57.25 minutes. As a first step, the process of focus groups foresees the activity of *sharing and comparing* (Nyumba et al., 2018). Participants can share their views on the subject and compare them. In addition, the interviewer keeps probing, searching for better and more complete descriptions of viewpoints as the participants describe their experiences (Nyumba et al., 2018; Morgan, 2021). During a 10-minute brainstorming session, participants created and then compared their digital sticky notes with those of the other participants. As a second step, we employed the activity *organize and conceptualize* to consolidate the generated information (Nyumba et al., 2018).

Group Nr.	Gender	Area of Expertise	Duration
Group 1	Female	Student, Industrial Engineering (M.Sc.)	61 minutes
	Female	Student, Global Business Management (B.Sc.)	
	Male	Student, Business Administration (M.Sc.)	
	Male	Student, Business Administration (B.Sc.)	

Group 2	Female Male	Research Assistant, Health Economics Research Assistant, Business Information Systems Engineering	42 minutes
Group 3	Female Female Male Male	Student, Industrial Engineering (B.Sc.) Student, Business Administration (B.Sc.) Student, Industrial Engineering (M.Sc.) Research Assistant, Business Information Systems Engineering	56 minutes
Group 4	Male Male Male	Student, Business Administration (M.Sc.) Founder, Software Engineer Employed, Customer Rollout Engineer	70 minutes

Table 1. Overview of Interviewees for the First Evaluation Cycle.

Second, we employed expert interviews to evaluate our feature set against the criteria of comprehensibility and level of detail. Expert interviews are appropriate for consulting persons with specific domain insights (Helfferich, 2014) and contributing as an evaluation element to the justification of a research artifact (Kaiser, 2014). We interviewed domain experts in a semi-structured nature to guarantee knowledge relevant to the subject matter and provide comparability of data quality between interviewees (Kaiser, 2014). Here, we also conducted the interviews online (i.e., Microsoft Teams and Mural) and recorded the sessions. On average, the interviews lasted 62 minutes. Before their interview, we shared the artifact with the experts to ensure sufficient time to apprehend the structural characteristics' names and feature names (including their developed definitions).

Group Nr.	Gender	Area of Expertise	Duration
Expert 1 (E1)	Female	Research Assistant, Digital Learning	61 minutes
Expert 2 (E2)	Male	Research Assistant, Finance and Accounting	60 minutes
Expert 3 (E3)	Female	Employee, Center for Learning and Teaching in Higher Education	65 minutes

Table 2. Overview of Interviewees for the Second Evaluation Cycle.

4 Results

The conversion of topics into features is initiated by identifying noun constellations, leading to the initial feature set (Figure 2). Topic 2, with 19.8%, and Topic 4, with 14.7% of all tokens, constitute the two largest topics. During the evaluation, we saw that they cover a considerable part of the features with "Course Elements and Activities" and "Teaching and Presentation Style" accounting for 16 features out of 25 features identified from the topic modeling. We then assess the additional data collected in the form of reports, blogs, or FAQ pages. After the evaluation, we stand with 32 features and seven structural characteristics. In the following we outline important comments from the focus group and expert interviews. Table 3 shows the identified features and their reference from topics or additional data.

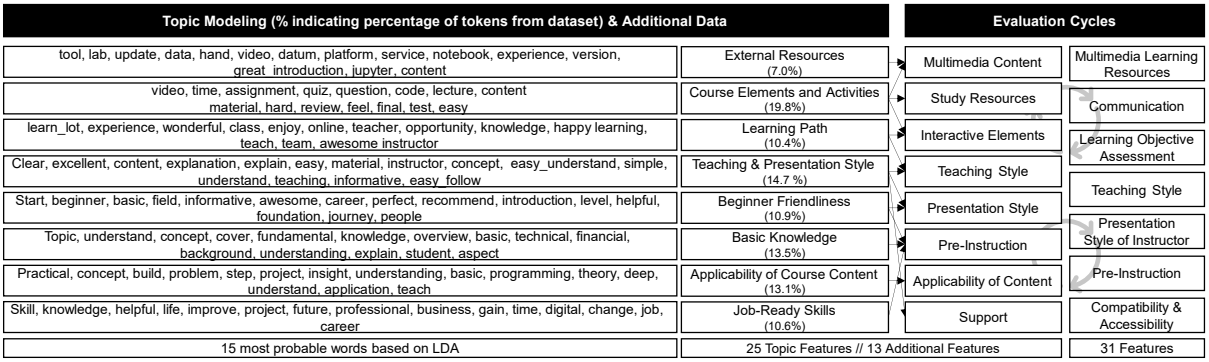

Figure 2. Results from Topic Modelling and Evaluations.

Nr.	Feature	Confirmed by							Origin
		G1	G4	G2	G3	E1	E2	E3	
(1) Multimedia Learning Resources						W	*	*	T1
#1	Video Content	B	B	B	B	R	*	*	T1, U3, S4, Y2, X4
#2	Audio Content	AD	B	B	AD	R	*	*	U3
#3	Primary Learning Resources	B	B	AD	B	W, R	*	*	T1, U3, S5, Y4, X3
#4	Additional Learning Resources	B	AD	AD	B	R	*	*	T1, T6, U3, S6, Y4, X3
#5	Face-to-Face Sessions	B	B	AD	B	R	*	*	C3
#6	Digital Consultation Session	⊘	⊘	⊘	⊘	NE	*	*	-
(2) Communication						*	*	*	T1, T2
#7	Discussion Forum	B	B	B	B	R	*	*	T1, U3, S1, Y1, X5
#8	Messaging/Chatting	B	B	B	B	R	*	*	C2, U3, Y7
#9	Automatic Messages	B	AD	AD	AD	R	*	*	U3
#10	Educational Announcements	⊘	AD	AD	B	R	*	*	U3, S1
(3) Learning Objective Assessment						*	*	*	T1, T2
Time of Assessment						NE	*	*	-
#11	Formative Assessments	B	B	AD	B	*	*	*	U5
#12	Summative Assessments	B	AD	AD	AD	*	*	*	U5
Type of Assessment						NE	*	*	-
#13	Assignments	B	AD	AD	AD	*	R	*	T1, U3,Y2, X2
#14	Practice Activities	B	AD	B	B	*	*	*	U5, S4, Y2, X6
Assessment Outcome						NE	*	*	-
#15	(Micro) Certificate	B	B	AD	B	R	*	*	C1, U3, Y3
#16	Learning Progress Representation	NE	B	B	B	*	*	*	
#17	Timely Learning Feedback	B	AD	B	AD	MO	W	W	T1
(4) Teaching Style						*	R	*	T2, T4
#18	Real World Examples/Application	B	AD	AD	AD	R	*	*	T2, U3, S4
#19	Problem Solution Suggestions	AD	AD	AD	AD	R	*	*	T4
#20	Domain-Relevant Tools	B	AD	AD	AD	MO	*	*	T8, S4, Y6, X9
(5) Presentation Style of Instructor						*	*	*	T2, T4, T5
#21	Speaking Pace and Tone	AD	B	AD	AD	*	*	*	T2
#22	Simple Language	AD	AD	AD	AD	*	*	*	T5, U4
#23	Gestures and Facial Expressions[1]	AD	AD	AD	AD	*	*	*	T2
#24	Instructors Engaging Demeanor	B	AD	AD	AD	*	*	*	T2
(6) Pre-Instruction						*	*	*	T2, T3, T5, T6, T7
#25	Actuality of Contents	AD	B	AD	AD	*	*	*	T2, T8
#26	Alignment of Content Display to Prior Knowledge	AD	B	B	B	W	*	*	T2
#27	Glossary	AD	AD	AD	AD	R	*	*	Y8
#28	Course Outline Planner	NE	AD	B	B	*	*	*	
(7) Compatibility & Accessibility						*	*	*	T8
#29	Subtitles	AD	B	AD	AD	*	*	*	C4, U1, S2, Y5, X4
#30	End Device Optimization	AD	AD	B	B	*	*	*	C5, U3, S3, X7
#31	Contents Adapted to Impairments	⊘	AD	AD	AD	W	R	*	C6, U4, X1
#32	Easy Login	⊘	NE	AD	AD	*	*	*	-

[1]Only applicable to "Video Content"; **B** = emerged during brainstorming; **AD** = emerged after discussion; **NE** = newly emerged during the interview

⊘ = did not appear or added later

NE = newly emerged during the interview; **W** = new wording of feature; **R** = rephrase definition; **MO** = moved to different structural characteristic; ***** = confirmed

T = from Topics
U = Udemy
S = Skillshare
U = Udacity
X = edX

Table 3. Overview of Confirmed and Evaluated Features.

In the first evaluation cycle, the initial feature set comprises the eight structural characteristics: (1) Multimedia Content, (2) Study Resources, (3) Interaction Elements, (4) Teaching Style, (5) Presentation Style of Instructor, (6) Pre-Instruction, (7) Applicability of Content and (8) Support based on the eight topics. The number of topics or now structural characteristics changed during the evaluation process as participants indicated that the structural characteristic "Multimedia Content" made sense to them. Originally, we subsumed the features #3 Primary Learning Resources and #4 Additional Learning Resources under this characteristic. However, the interviewees did not find the distinction intuitive. Thus, the structural characteristics "Multimedia Content" and "Study Resources" were united under the characteristic

"Multimedia Content". Furthermore, after the focus groups we added three features #16 Learning Progress Representation, #28 Course Outline Planer, and #32 Easy Login. Group one argued regarding #16 Learning Progress Representation that it appears motivating to see how far along students are in the overall progress of the course. At the same time, participants found it equally encouraging to be able to compare themselves with their peer group to challenge their learning progress against other learners. In connection to #28, participants mentioned from their own experience that an online course should offer an overview of dates and recommend learning times, which is helpful for self-organization. Another feature suggested by the second group is feature #32 Easy Login. The feature provides the ability to view and interact with individual course elements before enrollment. Overall, the structural characteristics of (2) Communication and (3) Learning Objective Assessment and their features were mentioned relatively often in all groups, especially in the brainstorming session, or were subsequently confirmed in the group discussion. In contrast, (4) Teaching Style and (5) Presentation Style of Instructor were hardly mentioned in the brainstorming phases, but the participants evaluated the features as useful and justified after the subsequent definition of the feature. Besides, one participant rightfully pointed out that feature #23 Gestures and Facial Expressions only applies to videos, and accordingly, we decided to include a respective note. In addition, we made a change to feature #31 Contents Adapted to Impairment after the first round. The feature was initially called "Variations in presentation", which was understood differently by the participants than intended. Participants assumed that the feature describes the different presentation of content or that one could navigate in the video. After rephrasing to "Contents Adapted to Visual Impairments", participants agreed to the new wording of feature #31. However, participants later on emphasized that this wording was too narrow from their viewpoint. They suggested rewording the feature to "Contents Adapted to Impairment" to include impairments in general. Finally, after the fourth round of interviews, we observed a high degree of alignment between the features we included in our set and the features that participants brainstormed and discussed during the focus group interviews. Accordingly, we ended the first evaluation cycle; the criterion of completeness was met.

In the second evaluation cycle, our feature set comprises 7 structural characteristics: (1) Multimedia Content, (2) Communication, (3) Learning Objective Assessment, (4) Teaching Style, (5) Presentation Style of Instructor, (6) Pre-Instruction, (7) Compatibility & Accessibility. After the expert interviews, we mapped features #17 Timely Learning Feedback and #20 Domain-Relevant Tools to other structural characteristics based on the feedback. In addition, the experts challenged the wording of features #3 Primary Learning Resources, #17 Timely Learning Feedback, #26 Alignment of Content Display to Prior Knowledge, and #31 Contents Adapted to Impairments, at which point new wordings emerged. E1 proposed reclassifying feature #17 Timely Learning Feedback from (4) Teaching Style to (3) Learning Objective Assessment. We agreed with the reasoning that feedback is closely related to learning objectives because the learning process only occurs when learners receive feedback. Later, E2 and E3 confirmed the reclassification. However, they stated that the original term "Feedback while Learning" is too broad (E2) and not understandable (E3). Respectively, E2 criticized the broadness since the feature's definition does not cover general feedback during a course but refers to the temporal proximity between the assessment and the response. Conversely, E3 criticized the comprehensibility of the wording, as the term is atypical from experience. No alternative term emerged during the interview; subsequently the feature was renamed #17 Timely Learning Feedback to account best for both comments. Moreover, E1 suggested a second reclassification for feature #20 Domain Relevant Tools. Specifically, from (1) Multimedia Content to (4) Teaching Style. Here, the argument is that the feature is related to #18 Real World Examples/Application and #19 Problem Solution Suggestions representing a concrete teaching element that the instructor employs. Later, E2 and E3 confirmed the reclassification.

During the expert interviews, we used the criteria of comprehensibility and level of detail to evaluate our feature set. Nonetheless, the new feature #6 Digital Consultation Session emerged that all experts mentioned independently of each other. According to E1, learning platforms increasingly implement this feature and provide it as an additional service in a course. The goal is to allow learners to book a one-on-one session with the instructor to answer questions. Consistent therewith, E2 and E3 described it as a synchronous interaction to put forward questions to the instructor to clarify aspects that learners may not understand on their own. Given that all experts independently identified it as a feature, we

decided to include it in the structural characteristic (1) Multimedia Content. After the third round of expert interviews, we established that the feedback was no longer related to basic comprehension questions, but experts mostly contributed formal comments (i.e., grammar or sentence structure). In addition, all experts confirmed the level of detail. Accordingly, we ended the second evaluation cycle because the criteria of comprehensibility and level of detail were met. Hereinafter, we present the final set of learning platform features that we were able to identify in our process. In total, 32 individual features are subsumed under seven structural characteristics. Additionally, each feature holds a definition that we deduced from the additional data obtained from the learning platform websites and the interviews.

The structural characteristic **Multimedia Learning Resources** comprises six features. First, *Video Content* refers to videos instructors record in advance, edit, and embed into the course to present information on a topic in a combination of visuals and audio. Learners can access and watch the video content (e.g., screencasts or animations) on-demand. Second, *Audio Content* describes a series of audio files (e.g., podcasts) that present the information on a topic to be learned via audio. Instructors record audio content in advance, edit them, and embed them into the course. Learners can access and listen to the audio content on-demand. Third, *Primary Learning Resources* contain the mandatory learning content. Usually, in the form of text, documents are embedded in the course. Usually, the resources are presentation slides used in the video content with continuous text or the transcript of the audio content, which are provided as PDF files or eBooks. Fourth, *Additional Learning Resources* are resources provided in addition to the Primary Learning Resources. These resources reinforce information conveyed and present information from a different perspective or show different application areas or use cases. Information from these resources is not relevant to learning objective assessments later in the course. The resources are embedded directly into the course or resemble a link collection. Sixth, *Face-to-Face Sessions* utilize technology (e.g., web conferencing) to allow learners to participate in a pre-scheduled one-to-many meeting with the course instructor. The session engages all learners in a conversation about the course content at a fixed time and for a fixed duration. Learners can participate and interact with third parties at their convenience. Seventh, *Digital Consultation Sessions* utilize technology (e.g., web conferencing) to allow the learner to schedule a one-on-one meeting with the course instructor. The goal of the session is to facilitate personalized scaffolding, meaning the instructor can break down the learning content into manageable chunks to help the learner overcome specific barriers to understanding the content.

The structural characteristic, **Communication** comprises four features. First, a *Discussion Forum* refers to the technical capability in a course whereby learners can raise questions to receive a response. The questions asked can be viewed and answered by the instructor and all enrolled learners. Questions are related to comprehension issues of the content within the topic. For example, when working on exercises, learners can ask technical questions concerning programming tools. Second, *Messaging/Chatting* refers to the capability to communicate electronically in real-time. Learners are enabled to contact and share information with learners and course instructors in private messages or in group chats. Furthermore, chatting allows to discuss factual information, for social interactions, tech-related questions, and clarify course scheduling. Third, *Automatic Messages* are messages sent autonomously through the learning platform to the learner (email or chat message). Messages are triggered by the learner's interaction with the course (e.g., welcome email after enrollment or successful completion of an assignment). They contain information on organizational aspects related to the course. Fourth, *Educational Announcements* are messages that learners receive from an instructor. They are used to communicate course updates or share free resources related to motivate topic relevance.

The structural characteristic, **Learning Objective Assessments** comprises seven features. First, *Formative Assessments* are assessments that take place during the course. Assessments are used after individual lessons allowing learners to assess how well they understand the course content. Learners can determine if there is a need to review previous lessons. A quiz at the end of a lecture is an example of a Formative Assessment. Second, *Summative Assessments* are assessments that take place at the end of a course. They are designed to assess how well learners understand the overall course content and whether learners have achieved the overall learning objectives of the course. A practical exam at the end of a course is an example of a Summative Assessment. Third, *Assignments* are assessments instructor's handout to learners that they must complete within a specified time. The assignment aims to help the learner

in critical thinking and achieve specific learning outcomes. After the assignment is completed, it is graded by the course instructor or other peers in the course. Fourth, *Practice activities* are guided experiences where learners can apply their knowledge. They are relevant and authentic to the course material and reflect real experiences and applications of the course topic. Interaction helps learners learn new content as they talk about it with someone else. Learners work through assignments in small groups to solve problems, develop critiques, and create artifacts. Fifth, *(Micro) Certificates* are an official attesting to a qualification successfully received by the learner. The certificate is provided upon successful completion of all required course elements, and the course enables learners to obtain a certificate from a renowned institution. Sixth, *Learning Progress Representation* describes the graphical visualization in percent or absolute numbers relative to the total number of learning units to be completed. The learner can check his progress and see other learners' progress. It is visualized in a dashboard design based on data retrieved from the online learning context. Seventh, *Timely Learning Feedback* refers to the feedback learners receive on their answers after completing a learning objective assessment. Suppose the answer deviates from the optimal solution. In that case, learners receive automated feedback through the platform, e.g., correct answers are stored for quizzes. Or humanoid feedback through the corrector, e.g., in the case of assignments, the score is explained to learners so they can better understand what the right answer would have been.

The structural characteristic, **Teaching Style**, comprises three features. First, *Real World Examples/ Applications* describe the instructors linking the imparted knowledge with the real world and show how the knowledge can be applied. For example, how clustering algorithms are used in customer management to identify and address customer groups with similar characteristics. Second, *Problem Solution Suggestions* refer to the instructor's activity of presenting schemes that enable the learner to decompose problems into their subproblems. This abstracts the complexity of the problem and enables a step-by-step approach to a problem. Such schemes can be used for example in debugging in programming. Third, *Domain-Relevant Tools* refer to tools relevant to the topic discussed within the course. Tools can be theoretical as well as technical. Theoretical tools are methods from the course's subject area used as a template when applying the knowledge (e.g., SWOT analysis). Technical tools include software (e.g., python, SPSS, or SAP) and enable the practical application of learned knowledge in the course. During the course, learners can access tools and apply them, for example, in exercises.

The structural characteristic, **Presentation Style of Instructor** comprises four features. First, *Speaking Pace and Tone* refer to the instructor's speech. The speech's speed is to be pleasantly fast so that it is possible to follow attentively. Besides, the lecturer's pronunciation is clear and important aspects are emphasized. Second, *Simple Language* is used to facilitate ease of understanding of the learning content and to support non-native speakers, idioms and nested sentence constructions are avoided. In addition, technical terms are refrained from being used without prior explanation, nor scientific language or slang is not employed. Learning content uses clear language and short, concise sentences. Third, *Gestures and Facial Expressions* refer to the posture and arm movements and facial expressions, eye contact, laughter, or positive expressions of instructors. This conveys a confident and calm appearance that allows the learner to follow the audio or video content. Fourth, *Instructors Engaging Demeanor* reflects the overall appearance of the instructor consisting of the features of speaking pace and tone, simpler language, gestures, and facial expressions that contribute to the perception of the lecturer as engaging.

The structural characteristic, **Pre-Instruction** comprises four features. First, *Actuality of Contents* refers to the period between the publication or updating of the course content and the time when the course content is viewed. Content loses its up-to-datedness when new types of information or new types of knowledge become available that partially or fully replace the former content. Second, *Alignment of Content Display to Prior Knowledge* The course progression or the presentation of individual learning contents is adapted to the prior knowledge of the learners. For this purpose, only learning content that is relevant for learners to achieve the learning objectives of the course is displayed in advance based on an initial test (e.g., quiz). Third, *Glossary* is a collection of domain-specific terminology that the instructor explains in the course. After the first explanation of a term, it is permanently available with explanations and translations enclosed. Usually reference is made to the course material where the terms are mentioned for the first time provided as a built-in tool or via the study resources. Fourth, *Course Outline*

Planner refers to a summary of the course topics. The outline includes the course description, number, title, topics, and course requirements. In addition, a schedule shows the most important dates, allows for scheduling exercise/learning times and exports the dates to the learners' calendar.

The structural characteristic, **Compatibility & Accessibility** comprises four features. First, *Subtitles* are lines of text that appear below video content. They provide written information about the content being shown. Continuous text fades in at the bottom to show spoken content. Subtitles can be translated into a foreign language to provide access to the learning content. This allows learners to visually perceive the audio associated with the video. Second, *End Device Optimization* refers to the technical capability of course content (multimedia content, resources, and interactive elements) to be accessed via different devices. Devices include laptops, tablets, and smartphones that learners can use as a means to consume and interact with the content. Third, *Contents Adapted to Impairments* characterizes the primary learning resource to be adapted to learners' impairments. For instance, for better visual comprehension and processing of course content, the resources provided should be adapted or resources should be specifically developed to help people with visual impairments (e.g., color blindness). For example, a larger font size, a different font type, and an adapted contrast and color design are helpful here. Fourth, *Easy Login* is a technical capability that allows learners to access a limited number of features or content in a course before or without registering. Learners can interact with the content to become familiar with it, making actual registration the next step.

5 Discussion

Platforms like Coursera, Udemy, Skillshare, Udacity, and edX have democratized knowledge access, fostering global educational goals (UN General Assembly, 2015). With over 220 million students worldwide (Class Central, 2023), these platforms are not just content repositories but complex ecosystems engaging both learners and instructors (Islind et al., 2021; Williamson, 2021). However, the proliferation of courses and content raises questions about learner satisfaction and the efficacy of platform features (Chowdhury et al., 2022; Faustmann et al., 2019).

Martin and Bolliger (2022) identify the different characteristics learner, course, instructor, and program and organization to look at learner satisfaction. In contrast, we identify seven characteristics that learners talk about (in course reviews and evaluations). Interestingly, the features belonging to the structural characteristics (4) Teaching Style and (5) Presentation Style of Instructor were hardly mentioned during our evaluation. During our third focus group interview, one of the participants remarked that the features of these structural characteristics were not considered because the participant took them for granted. This is in line with the viewpoint of Kano et al. 1984, who postulate that these features are seen as one-dimensional or must-be quality. Here, LDA shows the strength of analyzing the unobtrusive source of review data for user satisfaction (Jin et al., 2019). Furthermore, during the focus group and expert interviews, features were repeatedly brought up (i.e., digital learning assistants) that learners have not yet experienced themselves but would like to see implemented on learning platforms like MOOCs. The experts also emphasized that data-driven assistance systems will be used increasingly in the future to individualize learning paths, among other things, which current research strongly encourages (Zawacki-Richter et al., 2019; Sweller, 2020; Romero and Ventura, 2020). We do not see this as a limitation in our set of features, as the experts clearly confirm that the existing features reflect the status quo well on learning platforms. This is in line with our aim to enable exploitative innovation to make incremental improvements to existing products and services for the customer (i.e., learners) (Benner and Tushman, 2003; O'Reilly and Tushman, 2008; Božič and Dimovski, 2019).

Our theoretical implications are twofold. First, we provide researchers with a replicable methodological approach. While our approach differs from previous studies that focused on numeric and structured learning data (e.g., Hone and El Said (2016), Clow (2013)), we show how topic modeling helps to gain deep insights into unstructured textual data. Hence, our research adds to the literature by extracting latent topic dimensions that underlie unstructured online reviews. LDA enables us to offer a more detailed description of the main affordances that influence the learner's satisfaction. We hereby answer the calls of Li et al. (2020) to use NLP to create domain-specific knowledge and Reuver et al. (2018) to research

platforms in different industry settings. Moreover, our approach adds to the existing discourse that the use of data mining strategies is taking steps toward a level of productive maturity in education (Romero and Ventura, 2020). Second, our set of features extends an extensive body of research on understanding customer satisfaction on learning platforms (Jung et al., 2019; Hew et al., 2020; Anggraini et al., 2018; Rabin et al., 2019). Unlike existing research, we do not focus on a specific feature and do not rely on textual data alone but involve qualitative research and customers directly in the evaluation of features.

Our results also inform practice. First, for a digital platform to evolve successfully, complementors need to know how to rightly use and which platform features are offered by the platform core, otherwise knowledge boundaries may arise (Foerderer et al., 2019). Thus, the instructor as a complementor needs to be informed about which features are available to achieve learner satisfaction. Otherwise, missing knowledge might counteract the desired goal that the complementor uses the platform features in a value-adding manner (Darking et al., 2008; Gawer, 2014). The knowledge gaps ultimately lead to inconsistent actions and outcomes (Carlile, 2002; Foerderer et al., 2019). Second, our work provides an up-to-date overview of potential platform features. The features provide a roadmap for platform providers on which features are most valued by learners. This insight is invaluable for prioritizing development resources and ensuring that exploitative innovation efforts are focused on enhancing features that impact learner satisfaction. Furthermore, MOOC providers can use this data to refine their product offerings, ensuring that their platforms are not only user-friendly but also aligned with the educational goals and preferences of their target audience. The findings also inform curriculum designers and educators about features that resonate most with learners. This could lead to a more data-driven approach to curriculum development, focusing on integrating platform features that enhance learning efficacy.

6 Conclusion

The motivation of this paper was to develop domain-specific knowledge to provide platform features from learners' course reviews that platforms can use to improve learner satisfaction. At the same time, we provide a replicable methodological approach to identify learning platform features to answer our research question and analyze 142,976 learner reviews. The analysis through text mining and two evaluation cycles yielded a set of 32 noun-based product features and seven structural characteristics. Furthermore, the features provide a starting point to empirically determine the perception as attractive, one-dimensional, must-be, indifferent, or as reverse quality in the next step. Given the dynamic nature of technology and learner expectations, there's a rich opportunity for longitudinal studies. Future research could explore how the importance of various platform features evolves over time, providing insights into trends and shifts in learners' satisfaction. As with every research endeavor, our work is not without limitations and offers avenues for future research. The feature development involved subjective decisions during evaluations, interviewee assumptions or subjective ending conditions. However, we followed established research guidelines to minimize biases. Further, the collection and analysis of text data have their own challenges. We did not evaluate the performance differences of alternative preprocessing techniques or libraries on the topic modeling algorithm. Nonetheless, our research not only enriches the theoretical understanding of how unstructured data can provide deep insights into learner satisfaction but also demonstrates an approach that can be used for feature identification for exploitative innovation. By prioritizing knowledge dissemination among instructors on which features are available to them and focusing on developing features that match learner preferences, digital learning platforms can create a more engaging and satisfying learning environment. We lay the groundwork for future research, ensuring platforms are not just repositories of knowledge but thriving, learner-centric environments that foster inclusive, equitable, and quality education for all.

References

Abidin, M. (2021). "Stakeholders Evaluation on Educational Quality of Higher Education" *International Journal of Instruction* 14 (3), 287–308.

Alavi, M. and D. E. Leidner (2001). "Research Commentary: Technology-Mediated Learning—A Call for Greater Depth and Breadth of Research" *Information Systems Research* 12 (1), 1–10.

Alturki, A., G. G. Gable and W. Bandara (2011). "A Design Science Research Roadmap". In H. Jain, A. P. Sinha and P. Vitharana (eds.) *Service-Oriented Perspectives in Design Science Research*, pp. 107–123. Berlin, Heidelberg: Springer Berlin Heidelberg.

Anggraini, A., C. N. Tanuwijaya, T. Oktavia, M. Meyliana, H. Prabowo and S. Supangkat (2018). "Analyzing MOOC Features for Enhancing Students Learning Satisfaction" *Journal of Telecommunication, Electronic and Computer Engineering* (10), 67–71.

Bailom, F., H. J. Hinterhuber, K. Matzler and E. Sauerwein (1996). "Das Kano Model der Kundenzufriedenheit" *Marketing ZFP* 18 (2), 117–126.

Benner, M. J. and M. L. Tushman (2003). "Exploitation, Exploration, and Process Management: The Productivity Dilemma Revisited" *The Academy of Management Review* 28 (2), 238.

Berger, C., R. E. Blauth and D. Boger (1993). "KANO'S METHODS FOR UNDERSTANDING CUSTOMER-DEFINED QUALITY".

Berger, J., A. Humphreys, S. Ludwig, W. W. Moe, O. Netzer and D. A. Schweidel (2020). "Uniting the Tribes: Using Text for Marketing Insight" *Journal of Marketing* 84 (1), 1–25.

Bhattacherjee, A. (2001). "Understanding Information Systems Continuance: An Expectation-Confirmation Model" *MIS Quarterly* 25 (3), 351.

Blayone, T. J. B., R. vanOostveen, W. Barber, M. DiGiuseppe and E. Childs (2017). "Democratizing digital learning: theorizing the fully online learning community model" *International Journal of Educational Technology in Higher Education* 14 (1).

Blei, D. M., A. Ng and M. I. Jordan (2003). "Latent Dirichlet Allocation" *J. Mach. Learn. Res.* 3, 993–1022.

Božič, K. and V. Dimovski (2019). "Business intelligence and analytics use, innovation ambidexterity, and firm performance: A dynamic capabilities perspective" *The Journal of Strategic Information Systems* 28 (4), 101578.

Büschken, J. and G. M. Allenby (2016). "Sentence-Based Text Analysis for Customer Reviews" *Marketing Science* 35 (6), 953–975.

Carlile, P. R. (2002). "A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development" *Organization Science* 13 (4), 442–455.

Chang, V. (2016). "Review and discussion: E-learning for academia and industry" *International Journal of Information Management* 36 (3), 476–485.

Charmaz, K. (2006). *Constructing grounded theory. A practical guide through qualitative analysis.* 1. publ. London: SAGE Publ.

Chen, W.-K., J.-R. Chang, L.-S. Chen and R.-Y. Hsu (2022). "Using refined kano model and decision trees to discover learners' needs for teaching videos" *Multimedia tools and applications* 81 (6), 8317–8347.

Chowdhury, T. I., M. R. Hoque, P. Wanke, M. Z. Raihan and M. A. K. Azad (2022). "Antecedents of Perceived Service Quality of Online Education During a Pandemic: Configuration Analysis Based on Fuzzy-Set Qualitative Comparative Analysis" *Evaluation review* 46 (3), 235–265.

Chun Tie, Y., M. Birks and K. Francis (2019). "Grounded theory research: A design framework for novice researchers" *SAGE open medicine* 7, 2050312118822927.

Class Central (2023). *Massive List of MOOC Platforms Around the World in 2023. Comprehensive list of online course platforms worldwide, offering hundreds of thousands of courses.* URL: https://www.classcentral.com/report/mooc-platforms/ (visited on 05/03/2023).

Clow, D. (2013). "MOOCs and the funnel of participation". In: *Proceedings of the Third International Conference on Learning Analytics and Knowledge - LAK '13*. Ed. by D. Suthers, K. Verbert, E. Duval, X. Ochoa. New York, New York, USA: ACM Press, p. 185.

Conache, M., R. Dima and A. Mutu (2016). "A Comparative Analysis of MOOC (Massive Open Online Course) Platforms" *Informatica Economica* 20 (2/2016), 4–14.

Creswell, J. W. and J. D. Creswell (2018). *Research design. Qualitative, quantitative, and mixed methods approaches.* 5th edition, international student edition. Los Angeles, London, New Delhi, Singapore, Washington DC, Melbourne: SAGE.

Darking, M., E. A. Whitley and P. Dini (2008). "Governing diversity in the digital ecosystem" *Communications of the ACM* 51 (10), 137–140.

Debortoli, S., O. Müller, I. Junglas and J. vom Brocke (2016). "Text Mining for Information Systems Researchers: An Annotated Topic Modeling Tutorial" *Communications of the Association for Information Systems* 39, 110–135.

Dever, D. A., N. A. Sonnenfeld, M. D. Wiedbusch and R. Azevedo (2022). "Pedagogical Agent Support and Its Relationship to Learners' Self-regulated Learning Strategy Use with an Intelligent Tutoring System". In M. M. Rodrigo, N. Matsuda, A. I. Cristea and V. Dimitrova (eds.) *Artificial Intelligence in Education*, pp. 332–343. Cham: Springer International Publishing.

Faustmann, G., K. Kirchner, C. Lemke and D. Monett (2019). "WHICH FACTORS MAKE DIGITAL LEARNING PLATFORMS SUCCESSFUL?". In: *INTED2019 Proceedings*. Ed. by L. Gómez Chova, A. López Martínez, I. Candel Torres: IATED, pp. 6777–6786.

Foerderer, J., T. Kude, S. W. Schuetz and A. Heinzl (2019). "Knowledge boundaries in enterprise software platform development: Antecedents and consequences for platform governance" *Info Systems J* 29 (1), 119–144.

Garcia-Swartz, D. D., M. Muhamedagić and D. Saenz (2019). "The role of prices and network effects in the growth of the iPhone platform" *Technological Forecasting and Social Change* 147, 110–122.

Gawer, A. (2014). "Bridging differing perspectives on technological platforms: Toward an integrative framework" *Research Policy* 43 (7), 1239–1249.

Gimpel, H., D. Kleindienst, N. Nüske, D. Rau and F. Schmied (2018). "The upside of data privacy – delighting customers by implementing data privacy measures" *Electronic Markets* 28 (4), 437–452.

Gimpel, H., T. Manner-Romberg, F. Schmied and T. Winkler (2021). "Understanding the evaluation of mHealth app features based on a cross-country Kano analysis" *Electronic Markets* 31 (4), 765–794.

Griffin, A. and J. R. Hauser (1993). "The Voice of the Customer" *Marketing Science* 12 (1), 1–27.

Gupta, S. and R. Bostrom (2009). "Technology-Mediated Learning: A Comprehensive Theoretical Model" *Journal of the Association for Information Systems* 10 (9), 686–714.

Hacker, J., J. vom Brocke, J. Handali, M. Otto and J. Schneider (2020). "Virtually in this together – how web-conferencing systems enabled a new virtual togetherness during the COVID-19 crisis" *European Journal of Information Systems* 29 (5), 563–584.

Hein, A., M. Schreieck, T. Riasanow, D. S. Setzke, M. Wiesche, M. Böhm and H. Krcmar (2020). "Digital platform ecosystems" *Electronic Markets* 30 (1), 87–98.

Helfferich, C. (2014). "Leitfaden- und Experteninterviews". In N. Baur and J. Blasius (eds.) *Handbuch Methoden der empirischen Sozialforschung*, pp. 559–574. Wiesbaden: Springer Fachmedien Wiesbaden. URL: https://doi.org/10.1007/978-3-531-18939-0_39.

Hew, K. F. and W. S. Cheung (2014). "Students' and instructors' use of massive open online courses (MOOCs): Motivations and challenges" *Educational Research Review* 12, 45–58.

Hew, K. F., X. Hu, C. Qiao and Y. Tang (2020). "What predicts student satisfaction with MOOCs: A gradient boosting trees supervised machine learning and sentiment analysis approach" *Computers & Education* 145, 103724.

Hone, K. S. and G. R. El Said (2016). "Exploring the factors affecting MOOC retention: A survey study" *Computers & Education* 98, 157–168.

Hu, M. and B. Liu (2004). "Mining Opinion Features in Customer Reviews". In: *Proceedings of the 19th National Conference on Artifical Intelligence*: AAAI Press, pp. 755–760.

Impey, C. and M. Formanek (2021). "MOOCS and 100 Days of COVID: Enrollment surges in massive open online astronomy classes during the coronavirus pandemic" *Social sciences & humanities open* 4 (1), 100177.

Islam, A. N. (2012). "The Role of Perceived System Quality as Educators' Motivation to Continue E-learning System Use" *AIS Transactions on Human-Computer Interaction* 4 (1), 25–43.

Islind, A. S., L. Norström, H. Vallo Hult and S. R. Olsson (2021). "Socio-Technical Interplay in a Two-Sided Market: The Case of Learning Platforms". In C. Metallo, M. Ferrara, A. Lazazzara and S. Za (eds.) *Digital Transformation and Human Behavior*, pp. 33–53. Cham: Springer International Publishing.

Jansen, J. J. P., G. George, F. A. J. van den Bosch and H. W. Volberda (2008). "Senior Team Attributes and Organizational Ambidexterity: The Moderating Role of Transformational Leadership" *Journal of Management Studies* 45 (5), 982–1007.

Janson, A., M. Söllner and J. M. Leimeister (2020). "Ladders for Learning: Is Scaffolding the Key to Teaching Problem-Solving in Technology-Mediated Learning Contexts?" *Academy of Management Learning & Education* 19 (4), 439–468.

Jin, J., Y. Liu, P. Ji and C. K. Kwong (2019). "Review on Recent Advances in Information Mining From Big Consumer Opinion Data for Product Design" *Journal of Computing and Information Science in Engineering* 19 (1).

Jung, E., D. Kim, M. Yoon, S. Park and B. Oakley (2019). "The influence of instructional design on learner control, sense of achievement, and perceived effectiveness in a supersize MOOC course" *Computers & Education* 128, 377–388.

Kaiser, R. Dr. (2014). *Qualitative Experteninterviews*. *Konzeptionelle Grundlagen und praktische Durchführung.* Wiesbaden: Springer VS.

Kano, N. (2001). *Life Cycle and Creation of Attractive Quality.* Proceedings of the 4th International Quality Management and Organisational Development Conference (QMOD).

Kano, N., N. Seraku and F. Tsuji (1984). "Attractive quality and must-be quality" *he Journal of Japanese Society for Quality Control* (14:2), 39–48.

Li, Y., M. a. Thomas and D. Liu (2020). "From semantics to pragmatics: where IS can lead in Natural Language Processing (NLP) research" *European Journal of Information Systems* 30 (5), 569–590.

Löfgren, M., L. Witell and A. Gustafsson (2011). "Theory of attractive quality and life cycles of quality attributes" *The TQM Journal* 23 (2), 235–246.

March, S. T. and G. F. Smith (1995). "Design and natural science research on information technology" *Decision Support Systems* 15 (4), 251–266.

Martin, F. and D. U. Bolliger (2022). "Developing an online learner satisfaction framework in higher education through a systematic review of research" *International Journal of Educational Technology in Higher Education* 19 (1).

Matzler, K., H. H. Hinterhuber, F. Bailom and E. Sauerwein (1996). "How to delight your customers" *Journal of Product & Brand Management* 5 (2), 6–18.

Mayer, R. E., L. Fiorella and A. Stull (2020). "Five ways to increase the effectiveness of instructional video" *Educational Technology Research and Development* 68 (3), 837–852.

Merton, R. K. (1987). "The Focussed Interview and Focus Groups: Continuities and Discontinuities" *Public Opinion Quarterly* 51 (4), 550.

Mimno, D., H. M. Wallach, E. Talley, M. Leenders and A. McCallum (2011). "Optimizing Semantic Coherence in Topic Models". In: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. USA: Association for Computational Linguistics, pp. 262–272.

Mkpojiogu, E. O. C. and N. L. Hashim (2016). "Understanding the relationship between Kano model's customer satisfaction scores and self-stated requirements importance" *SpringerPlus* 5, 197.

Morgan, D. L. (2021). "Robert Merton and the History of Focus Groups: Standing on the Shoulders of a Giant?" *The American Sociologist*.

Müller, O., I. Junglas, J. vom Brocke and S. Debortoli (2016). "Utilizing big data analytics for information systems research: challenges, promises and guidelines" *European Journal of Information Systems* 25 (4), 289–302.

Myers, M. D. and M. Newman (2007). "The qualitative interview in IS research: Examining the craft" *Information and Organization* 17 (1), 2–26.

Nickerson, R. C., U. Varshney and J. Muntermann (2013). "A method for taxonomy development and its application in information systems" *European Journal of Information Systems* 22 (3), 336–359.

Nili, A., M. Tate and D. Johnstone (2017). "A Framework and Approach for Analysis of Focus Group Data in Information Systems Research" *Communications of the Association for Information Systems* 40, 1–21.

Nyumba, T. O., K. Wilson, C. J. Derrick and N. Mukherjee (2018). "The use of focus group discussion methodology: Insights from two decades of application in conservation" *Methods in Ecology and Evolution* 9 (1), 20–32.

O'Reilly, C. A. and M. L. Tushman (2008). "Ambidexterity as a dynamic capability: Resolving the innovator's dilemma" *Research in Organizational Behavior* 28, 185–206.
Papp, R. (2000). "Critical Success Factors for Distance Learning" *Americas Conference on Information Systems (AMCIS)*, 1858–1861.
Park, C. W., S. Milberg and R. Lawson (1991). "Evaluation of Brand Extensions: The Role of Product Feature Similarity and Brand Concept Consistency" *Journal of Consumer Research* 18 (2), 185.
Peng, X. and Q. Xu (2020). "Investigating learners' behaviors and discourse content in MOOC course reviews" *Computers & Education* 143, 103673.
Rabin, E., Y. M. Kalman and M. Kalz (2019). "An empirical investigation of the antecedents of learner-centered outcome measures in MOOCs" *International Journal of Educational Technology in Higher Education* 16 (1).
Reuver, M. de, C. Sørensen and R. C. Basole (2018). "The Digital Platform: A Research Agenda" *Journal of Information Technology* 33 (2), 124–135.
Rice, K. L. (2006). "A Comprehensive Look at Distance Education in the K–12 Context" *Journal of Research on Technology in Education* 38 (4), 425–448.
Romero, C. and S. Ventura (2020). "Educational data mining and learning analytics: An updated survey" *WIREs Data Mining and Knowledge Discovery* 10 (3).
Schmiedel, T., O. Müller and J. vom Brocke (2019). "Topic Modeling as a Strategy of Inquiry in Organizational Research: A Tutorial With an Application Example on Organizational Culture" *Organizational Research Methods* 22 (4), 941–968.
Sonnenberg, C. and J. vom Brocke (2012). "Evaluation Patterns for Design Science Research Artefacts". In: *Practical Aspects of Design Science*. Ed. by M. Helfert, B. Donnellan. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 71–83.
Sun, R. and S. Gregor (2023). "Reconceptualizing platforms in information systems research through the lens of service-dominant logic" *The Journal of Strategic Information Systems* 32 (3), 101791.
Sweller, J. (2020). "Cognitive load theory and educational technology" *Educational Technology Research and Development* 68 (1), 1–16.
Turney, P. D. and P. Pantel (2010). "From Frequency to Meaning: Vector Space Models of Semantics" *Journal of Artificial Intelligence Research* 37, 141–188.
UN General Assembly (2015). *Transforming our world : the 2030 Agenda for Sustainable Development*. URL: https://www.refworld.org/docid/57b6e3e44.html (visited on 11/10/2023).
Wambsganß, T., A. Schmitt, T. Mahnig, A. Ott, N. Ngo, J. Geyer-Klingeberg, J. Nakladal and J. M. Leimeister (2021). *The Potential of Technology-Mediated Learning Processes: A Taxonomy and Research Agenda for Educational Process Mining*.
Williamson, B. (2021). "Making markets through digital platforms: Pearson, edu-business, and the (e)valuation of higher education" *Critical Studies in Education* 62 (1), 50–66.
Witell, L., M. Löfgren and J. J. Dahlgaard (2013). "Theory of attractive quality and the Kano methodology – the past, the present, and the future" *Total Quality Management & Business Excellence* 24 (11-12), 1241–1252.
Yang, B., Y. Liu, Y. Liang and M. Tang (2019). "Exploiting user experience from online customer reviews for product design" *International Journal of Information Management* 46, 173–186.
Yu, Q. (2022). "Factors Influencing Online Learning Satisfaction" *Frontiers in psychology* 13, 852360.
Zawacki-Richter, O., V. I. Marín, M. Bond and F. Gouverneur (2019). "Systematic review of research on artificial intelligence applications in higher education – where are the educators?" *International Journal of Educational Technology in Higher Education* 16 (1).