

ON-DEMAND VIDEO TAGGING, ANNOTATION, AND SEGMENTATION IN LECTURE RECORDINGS TO ENHANCE E-LEARNING EFFECTIVENESS

Ken D. Nguyen and Muhammad Asadur Rahman

Department Dept of Computer Science & Information Technology,
Clayton State University, Georgia, USA

ABSTRACT

The COVID-19 pandemic has forced much of the academic world to transition into online operations and online learning. Interactions between the teachers and students are carried out via online video conferencing software where possible. All video conferencing software available today is designed for general usage and not for classroom teaching and learning. In this study, we analyzed the features and effectiveness of more than a dozen major video conferencing software that are being used to replace the physical face-to-face learning experiences. While some of the video conferencing software has pause feature but none allow annotation and segmentation of the recording. We propose tagging and annotation during the live streaming to improve direct access to any portion of the recorded video. We also propose automatic segmentation of the video based on the tagging so that the video is short, targeted, and can easily be identified.

KEYWORDS

Live and On-Demand Video, Video Tagging, Video Annotation, Video Segmentation, E-Learning, Remote Learning, Synchronous Lectures, User Interface Design, Human-Computer Interaction.

1. INTRODUCTION

Since the outbreak of the COVID-19 pandemic, the ways in which we teach and learn have shifted dramatically from largely in-person to completely virtual. Teaching in the fields of computer science and technology often requires close interaction between students and instructors, and online teaching and learning require tremendous organization and discipline. Furthermore, effective online education greatly depends on the vehicle in which a course is delivered. Software tools that are used in synchronous or asynchronous delivery of courses must have user-friendly interfaces to promote effective and timely interaction between instructors and students.

The COVID-19 pandemic has shifted much of the US education system onto remote learning, changing from instructor-led face-to-face interactive mode to distance learning via online web learning management systems (LMS) such as Desire2Learn, KMI LMS, GoToTraining, Grovo, Edwiser remUI, Moodle, Edmego, SmarterU, Paradiso Solutions, DHx Software, MindQuest Learning, Brainier, Talent LMS [1,2,3,4,5,6,7,8,9,12,13,14] ScholarLMS, Eurekos LMS, InnTier, Knolyx, Wisetail, Learnupon, GO[15], etc.

Video lectures and synchronous video conferencing are two primary mechanisms for instructors to emulate the in person native interactive environment of teaching and learning. A recent study [16] conducted on a random control monitoring of 157 students in a for-credit online course at a 4-year university found that on average, students watch all assigned and required videos with embedded assessment 1.29 times, about 9% of videos are skipped by students, and about 2% of the videos are watched after the assigned deadline. These figures indicate that students mostly watch the videos once to complete the assignment to get the grade and then almost completely ignore the video. This intriguing fact seems to contradict the common assumption that a recorded lecture would allow the student to review the material multiple times which could lead to better understanding of course materials and enhance their performance. In addition, in the studies [17,18], recorded lectures seem to support significant improvement in student performance if they are used as supplements.

Studies have shown that segmenting videos of 6 minutes or less in length is more effective for leaning [17,18,19,20], It allows learners to engage with small pieces of new information, gives them control over the flow of information, and helps encode the information into long-term memory. Long videos are ineffective as they overload the intrinsic cognitive capability of learning working memory [19,20,21,22,23].

While remote learning is not new, effective remote learning has always been a challenge as most younger and inexperienced learners need constant guidance; moreover, courses with higher levels of abstraction and logical thinking such as in the STEM (Science, Technology, Engineering, Mathematics) fields require repetitive guidance and often hands-on experience, thus requiring instructor-led class time as seen in most education institutions in the US and around the world. Because on-site operations at educational institutions in the US are now closed to slow down the spread of the coronavirus, most instructors depend on video conferencing tools such as Zoom, Microsoft Teams, WebEx, GoogleMeet, GoToMeeting, TeamViewer, YouTube, Facebook, Tiktok, Telegram, Slack, Discord and the likes to replicate direct hands-on face-to-face guidance during remote learning. This switch is a big challenge in teaching as instructors are often not well-equipped or trained, nor do they have the time to become effective video content makers and video editors. Moreover, none of the listed video tools are capable of or designed for handling video conferencing in academic environment.

Table 1. Supported features by various online synchronous video meeting software.

Software	Recording	Pause/Resume	Annotated	Video Segmentation	Transcription	Chat	Multi-User
Microsoft Teams	Yes	No	No	No	Yes	Yes	Yes
Zoom	Yes	Yes	No	No	Yes	Yes	Yes
Google Meet	Yes	No	No	No	Yes	Yes	Yes
WebEx	Yes	Yes	No	No	Yes	Yes	Yes
Team Viewer	Yes	Yes	No	No	No	Yes	Limited
Telegram	Limited	No	No	No	No	Yes	Limited
Facebook Live Stream	Yes	Yes	No	No	No	Yes	Limited
Goto Meeting	Yes	No	No	No	Yes	Yes	Yes
Discord	No	No	No	No	No	Yes	Limited
Slack	No	N/A	No	No	No	Yes	Limited
Youtube Live	Yes	No	Yes	No	Yes	Yes	No
Tiktok	No	No	No	No	No	No	No

Table 1 depicts a summary of our review of popular software that support online video recording and posting. We found that none of the popular video conferencing tools being used in education have proper interface and functionality to allow the instructors effectively create segmented instructional videos for use in their courses. From our study, we propose new feature and interface design to provide the much needed and effective tool for instructors to create targeted and topic specific videos to help promote effective remote teaching and learning.

2. RECORDING TAGGING AND ANNOTATING PROPOSAL

Synchronous video class meetings between faculty and students tend to cover many topics that span over an hour or longer. The recordings of these meetings could be more effectively used if they could be fragmented into individual segments such that each segment is focused on a specific item, topic, question, or idea. There should be an easy option for quickly marking the starting and ending of each fragment on the fly during class discussion. The author should be able to go back to the recorded video to adjust the markers and then extract the fragments as smaller videos for posting.

2.1. New Features to Consider

In our study, we look at Microsoft Teams, Zoom, WebEx, and Google Meets as the most popular conferencing software being used in academia because they can handle many users simultaneously and can integrate with institutional digital infrastructure. These software allow the user to trim the recordings and have the recordings transcribed. They also have a good auto-generated transcript service for recording videos. Zoom and WebEx allow the recording to be paused and resumed during the meeting, thus allowing some of the discussion to be included in the recordings. While these software have many great features, the final product – the recorded meeting videos – are often long and not effective for learning in academia. While the instructor can prepare a pre-recorded video using a camcorder or a recording software such as Kaltura Capture[24] or similar software, these videos do not represent the class dynamic or the specific topics and concepts in a way that the students would approach in real time. In addition, the majority of course-specific materials are specific to individuals, groups of students, and each semester. This issue is intensified much more in courses that have frequent updates to materials such as IT and CS. Therefore, synchronous meetings and recorded meeting videos are still very popular. Instructors are often faced with the decision to edit and fragment videos or leave them as-is. The second option is often the choice as time, resources, and training are limited.

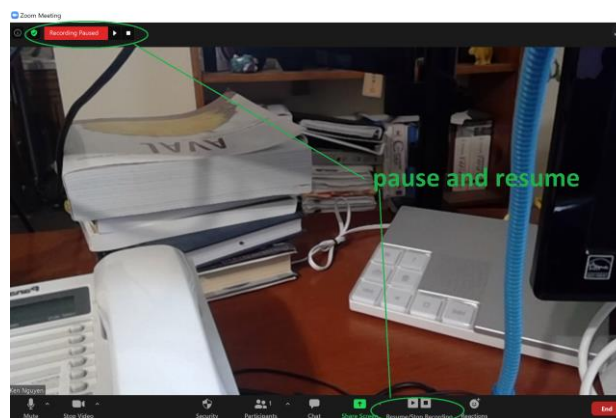


Figure 1. Zoom meeting with Pause/Resume option

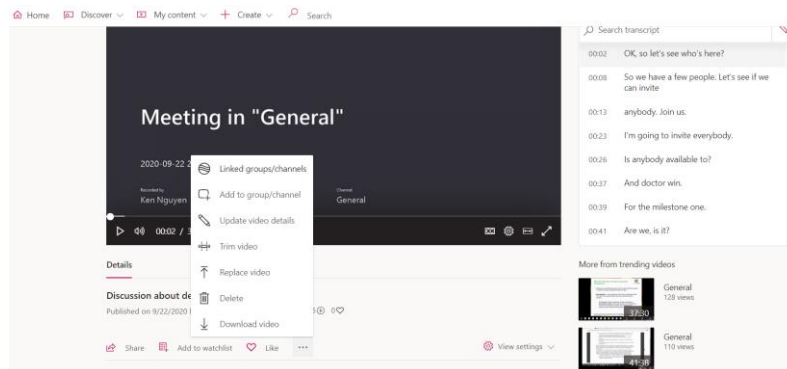


Figure 2. Microsoft Teams Streaming Transcription

Moreover, video editing software such as Adobe Premiere and others require tremendous amounts of time and effort, and video editing and video production are skills in which most instructors are not trained or exposed to. Figures 1 shows a screen capture of Zoom synchronous video meeting and Figure 2 shows the transcription of a Microsoft Teams recording.

Moreover, the option to pause the recording has a big impact as to which institutions can and cannot use the software. Without the pause option, the recording will be contiguous, thus making the recording more transparent and easier to verify any possible allegation that arises from attendees. Not allowing pausing on the recording would allow tracing and tracking of any potential legal claim against the meeting participants. Any editing of the original recording would show the alteration, doctoring, or tampering of the evidence. Therefore, software such as Microsoft Teams are the primary choice for most public institutions. Not allowing pausing on the recording would allow tracing and tracking of any potential legal claim against the meeting participants. On the other hand, having the pause option allows certain parts of the meeting to be excluded from the recording; thus, this option works well for private and nonessential meetings. However, in almost all instructional synchronous meetings, there are many times a topic is reiterated, so the same questions may be asked by different individuals in the same meetings, and some repetitive and insignificant discussions are recorded during lectures. Thus, having all of these in the recording videos makes them very long, and students lose focus and interest. We want to have a feature that can automate the fragmentation process and retain the best of these two models.

2.2. Proposal to Enhance Video Recording with Targeted Tagging and Annotations

In terms of interrupting a recording, there could be two models of video conference recording: (a) those with a pause option such as Zoom and WebEx, and (b) those without a pause option. The platforms with the pause option allow the user to skip recording of certain periods of time when the discussion is not directly related to or of particular significance for those who may review the video later. These tools are good for personal and private environments, where there is less scrutiny and little chance that legal action may arise or be taken for or against what is being recorded. The second model, without the pause option, is often appropriate for public use and institutions involving many stakeholders. Thus, the recording should be continuous and untampered so that it can withstand all scrutiny, showing transparency, and can be used for legal purposes.

2.2.1. Designs

While these video conferencing applications are excellent at facilitating synchronous meetings, they lack certain features to support an effective online learning environment. For example, students generally lose focus when the video duration exceeds 6 minutes; that amount of time or less in length is more effective for learning[19]. Thus, having long recording sessions usually discourages students from reviewing the videos. Secondly, long recordings generally cover a lot of material and contain many interweaving concepts. It is best to make short video segments for each concept paired with illustrative examples. However, video making requires extensive amounts of preparation time and video editing skills that most faculty would not have, especially for those who have been teaching face-to-face and now are suddenly thrown into fully online instruction.

Our proposed design enhances these video conferencing applications so that the user can make annotations during the recording of conference meetings and have the recordings segmented into shorter videos automatically. The user then can make these video into shorter segments with specific context to their students. For video conference recording software that do not have the “pause” feature, we propose the recording enhancement as seen in Figure 3.

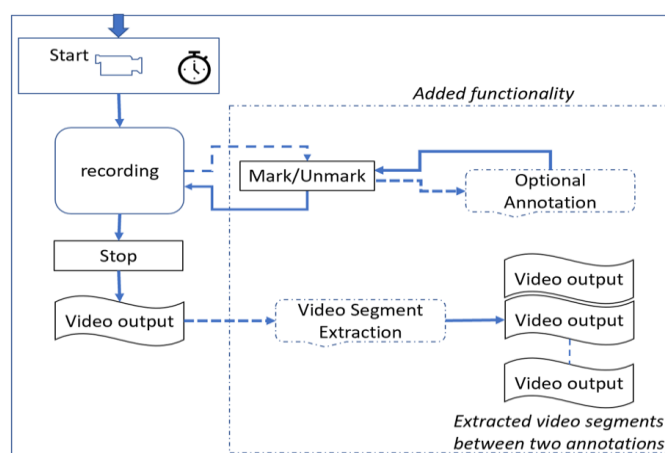


Figure 3. Enhanced recording feature for video conference software without a “pause” feature

The recording can run continuously from the moment it is started until the stop option is triggered. However, in our model, a permanent alternating “mark” and “unmark” button will be displayed on the interface, allowing the user to trigger segmentation markings on the recording. The button will alternate between mark and unmark. This option allows the user to place an annotation mark on the video at the moment of choice. A semi-transparent fading balloon for annotation will appear for a few seconds following the “mark,” allowing the user to type in a short title or description for the marked segment. If the user ignores the balloon, the first transcribed sentence immediately following the marking will be used to annotate the title of the recorded segment. This feature allows the user to keep full attention on video conferencing without distraction. The user can edit the automated text after the video recording session is finished.

For video conference recording software that do not have a “pause” feature, we propose the recording enhancement as seen in Figure 4.

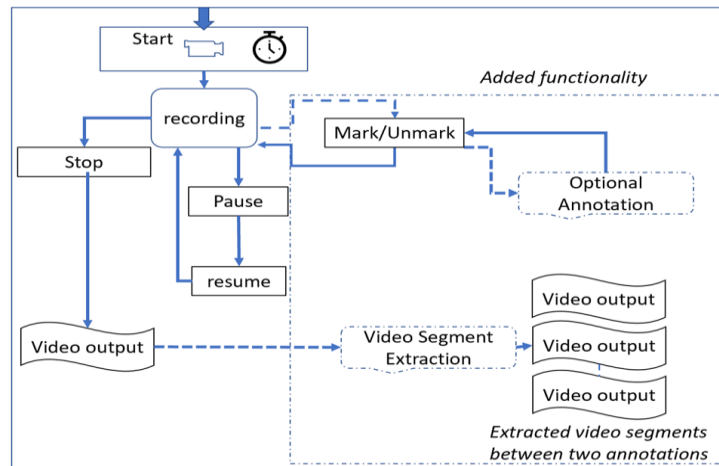


Figure 4: Enhanced recording feature for video conference with 'pause' feature

2.2.2. Recording with Enhanced Video Tagging for Marking and Annotating Segments of Recording Video

Enhanced tagging for the marking/unmarking and annotating feature would run in conjunction with the traditional “pause” option in any video recording service, where the user can put the recording on hold and resume the recording when ready. The pause and resume functions could share the same button with alternating options displayed. However, the user will be presented with a semi-transparent fading pop-up balloon/screen which can be annotated at the beginning of a recording segment. The user can either choose to click on the balloon and enter some text to annotate or describe what would be immediately discussed in the video right after the resume option is triggered or to ignore it completely. The balloon will fade away in a few seconds if the user is not clicking on it; however, an annotation button will appear next to the resume button so that the user can pull up the annotation text box.

Since the recording can be paused, the user would have full control of the recording. Thus, it is not necessary to add a new button that takes up precious interface space, especially when the user is using small mobile devices such as a smart phone. For this model, the pause button would function as a traditional pausing option. The enhancement appears when the user clicks on the annotated balloon and a message can be added. We also propose to have an “annotated” button to be displayed on the control bar which the balloon fades into. This button allows the user to bring back the annotation balloon to annotate the beginning of a recording segment.

2.2.3. Automatic Video Segmentation and Extraction with the New Tagging and Annotation Feature

This option is added into both models, giving the user an option to extract video recording segments between annotations as individual videos. In online education, this feature is a great time saving tool for preparing teaching materials and learning. An example use case would be as follows: a computer science professor describes the concept of a data structure, its usage, and how to implement it to students. Thus, he would have examples to illustrate the concept. The examples would be very appropriate and precise due to his experience and knowledge in the field. However, these kinds of examples may not relate well with students as they do not have the same background as the professor, and the precision of the examples may require a certain level of influence, ability to recall, utilize and readily apply previously learned concepts. Thus, when the user makes his own video lecture, the students may still not be able to comprehend it very

clearly. A discussion is usually followed by the posting of the lecture so students can ask for clarification or alternate explanations. These kinds of questions are individual and specific to help students to grasp the concept. Thus, the professor then responds to the students’ questions with specific examples and illustrations. These answers are often more related to the students. Thus, having a way to extract these specific illustrations and place them along with the lecturing concept will help the students to learn and allow the professor to utilize their time effectively. This use case assumes that the professor is given a time slot to interact with students. With current video conference tools, the professor does not have that option. He may have to make several videos for each of these examples. Certainly, these video segments will not be professionally done; however, most professors are not professional video editors anyway.

2.2.4. Legal Implication of Contiguous Recordings

The option to record a meeting without a pause option allows institutions and corporations to have a continuous and unaltered record that can withstand potential legal tests when there are complaints about what has been said or shared during the meeting. With our proposal, the original recordings are the same as the original system would produce, and our proposal allows the system to generate a by-product that is beneficial and supportive to an online academic environment.

2.3. Implementing the Enhanced Tagging and Annotation Features

To implement these proposed features with minimal modification to the system and without changing its structure, we propose adding only a button onto the interface to provide the user with a way to manually trigger the new feature. Figure 5 depicts the implementation of the design where an auxiliary file is being used to record all annotations during the recording session in sync with the recording clock.

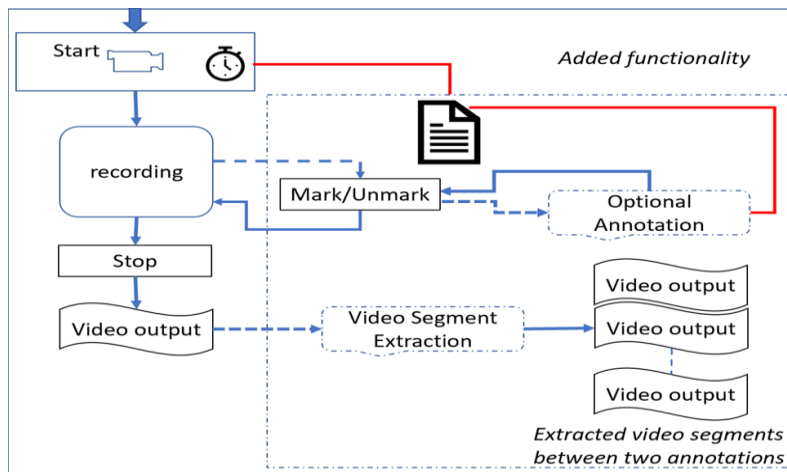


Figure 5: Enhanced recording feature for video conference with 'pause' feature

All annotations can be done as a floating semi-transparent text box interface. To support the creation of the video segments, we will rely on the timer and a description file. In general, most systems employ a clock to show elapsed time in the recording, which is also the length of the recording, and we will use the same clock to support the new feature. When the new button is triggered, the system recording duration time is saved. Similarly, every interaction with the new feature will be recorded with the content given, and the time relative to the video recording. For example, two minutes into the recording, the user presses the “mark/unmark” button, annotates a

message 3 seconds later, and then clicks the “pause/unmark” button 5 minutes after the start of the annotation. This allows the system to generate a 5 minute and 3 second long recording segment starting 2 minutes from the beginning of the recording. An example entry for the description file is:

```
2:00    2:03    the message    7:03
```

The message can either be superimposed on the beginning of the segment or kept as the beginning of the segment transcription.

Multiple annotations: To handle multiple annotations on the same segment, the user can always annotate the video by typing into the semi-transparent text box and hit “enter” to save the message. The system records the time when the first character is entered into the text box and the message for the future display in the video segment. For example, at 2 minutes into the recording, the user clicks the mark/unmark button, then types in a message 5 seconds later; then the user types another message 1 minute later and clicks the unmark button at 9 minutes into the recording. The system will record the following:

```
2:00    2:05    the first message    3:05    the second message    9:00
```

Thus, with this technique, the user can annotate any number of messages, each value is separated by a tab, and the first and last value of the entry indicates the time interval from which the original recording will be copied to create a new recording segment.

When the recording is stopped, the system will use these entries in the description file to generate the recording segments from the original recorded video, and the messages can be superimposed on each segment individually at the corresponding recorded time period.

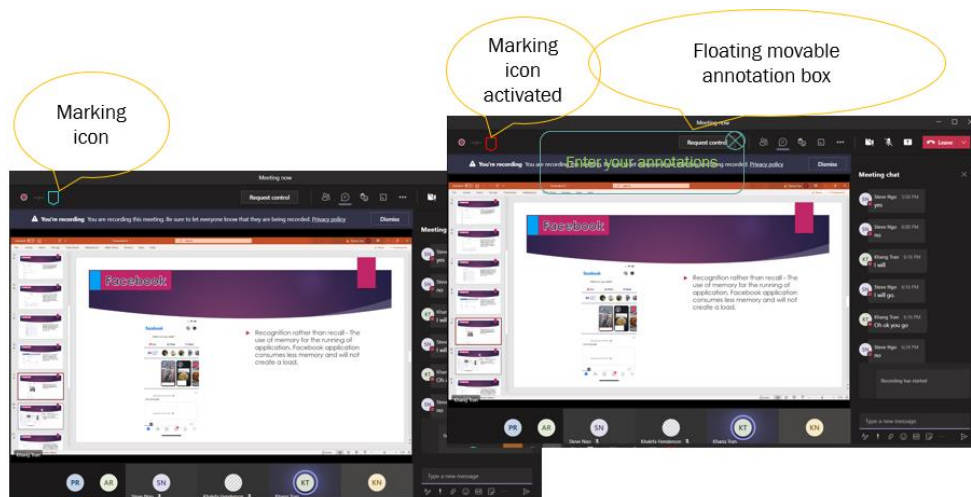


Figure 6: Interface for segment marking and annotation features

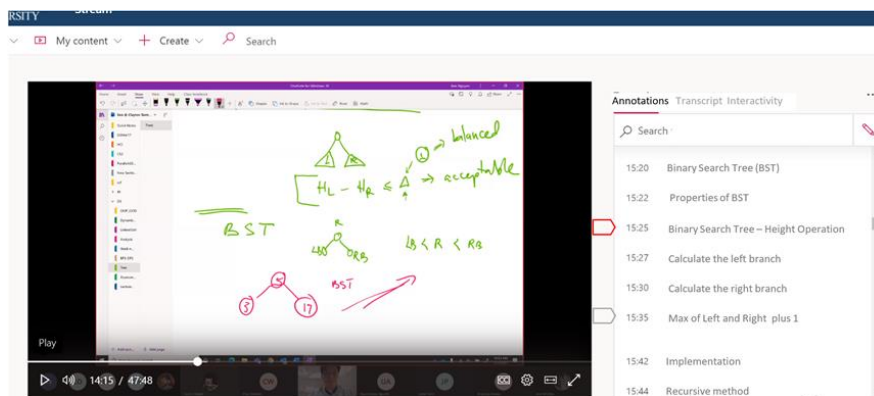


Figure 7. Segment Marking and annotations on Recording Videos

The user interface design for the segmentation marking and annotation features of the prototype system is given in Figure 6. The marking icon when clicked is activated which would become red to indicate the start of the segment. Clicking it again would mark the end of the segment. The annotation box is a floating box where annotation can be written. Figure 7 shows active marker and annotations on a recording.

3. CONCLUSIONS

In this study we analyzed over a dozen synchronous video meetings and conferencing software. We looked at the video delivery user interfaces that would be most appropriate and effective for student learning and engagement. From our study and recent research, we concluded that none of these software has the video features necessary for effective remote learning. We proposed features that would seamlessly integrate with any video conferencing software allowing the user to annotate the recording of the meeting on the fly, on-demand and obtain automated segmentation of the video. The segmentation feature would allow topic centric videos to be generated automatically from the recording without requiring further editing and minimal intervention. These new features can be added on to the existing video conferencing software without changing the underlying infrastructure. As future research, we would like to analyze and compare the effectiveness of these auto-generated targeted video segments against specialized video content created for similar topics.

REFERENCES

- [1] Tovutilms. "Learning management System". <https://www.tovutilms.com/> Date accessed January 28, 2021.
- [2] ELucid. "Build a Solid Corporate Training Culture with ELucid"., <https://elucidlearning.co/>. Accessed 26 Jan. 2021.
- [3] "KMI Learning". Custom ELearning Solutions., <https://www.kmlearning.com/>. Accessed 26 Mar. 2021.
- [4] GoToMeeting. "Online Meeting Software, Video Conferencing & Web Conferencing". <https://www.gotomeeting.com/>. Accessed 10 Dec. 2020.
- [5] Grovo. "The Microlearning Company.", 20 Dec. 2020, <https://www.grovo.com>
- [6] Edwiser. "LMS Solutions Under One Roof!". <https://edwiser.org/>. Accessed 5 Nov. 2020.
- [7] Moodle. "Open-Source Learning Platform". <https://moodle.org/>. Accessed 01 Oct. 2020.
- [8] Edmego Learning. "Edmego Learning Management System., 11 Mar. 2016, <http://www.edmegolearning.com/>. Accessed 18 Oct. 2020.
- [9] SmarterU. "Learning Management System | LMS". <https://www.smarteru.com/>. Accessed 12 Feb 2021.
- [10] Paradiso LMS. "Best Learning Management System (LMS) Software". <https://www.paradisosolutions.com/>. Accessed 30 Jan. 2021.

- [11] DHx Software. “Digital Learning Management System (LMS)”. <https://www.dhxsoftware.com/elearning>. Accessed 7 Feb. 2021.
- [12] LMS.Org. “Learning Management System Reviews”. <https://www.lms.org/>. Accessed 2 Mar. 2021.
- [13] TalentLMS. “Cloud LMS Software”. <https://www.talentlms.com/>. Accessed 6 Feb. 2021.
- [14] Thought Industries. “Powering the Business of Learning”. <https://www.thoughtindustries.com>. Accessed 15 Jan. 2021.
- [15] Complete List of Learning Management System Reviews, <https://www.lms.org/learning-management-systems-list/>. Accessed 23 Sep. 2021
- [16] Baker, R., Evans, B., Li, Q. et al. (2019) “Does Inducing Students to Schedule Lecture Watching in Online Classes Improve Their Academic Performance? An Experimental Analysis of a Time Management Intervention” *Res High Educ* 60, 521–552 <https://doi.org/10.1007/s11162-018-9521-3>
- [17] E. Péter and H. Ferenc (2017) "Analysis of video views in online courses" *40th International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia*, pp. 778-782, doi: 10.23919/MIPRO.2017.7973527.
- [18] Guo, P. J., Kim, J., & Rubin, R.(2014) “How video production affects student engagement: an empirical study of MOOC videos” *AMC Conference on Learning @ Scale*. Atlanta, GA.
- [19] Brame CJ. (2016) “Effective Educational Videos: Principles and Guidelines for Maximizing Student Learning from Video Content” *CBE Life Sci Educ.*;15(4):es6. doi:10.1187/cbe.16-03-0125
- [20] Birgili B, Seggie FN, Oğuz E, (2021) “The trends and outcomes of flipped learning research between 2012 and 2018: A descriptive content analysis” *Journal of Computers in Education*.1-30. doi:10.1007/s40692-021-00183-y
- [21] Suppan M, Stuby L, Carrera E, et al. (2021) “Asynchronous Distance Learning of the National Institutes of Health Stroke Scale During the COVID-19 Pandemic (E-Learning vs Video): Randomized Controlled Trial”. *J Med Internet Res*. 23(1):e23594. Published 2021 Jan 15. doi:10.2196/23594
- [22] Choe RC, Scurig Z, Eshkol E, et al.(2019) “Student Satisfaction and Learning Outcomes in Asynchronous Online Lecture Videos” *CBE Life Sci Educ*. 18(4):ar55. doi:10.1187/cbe.18-08-0171
- [23] Cooper KM, Ding L, Stephens MD, Chi MTH, Brownell SE.(2018) “A Course-Embedded Comparison of Instructor-Generated Videos of Either an Instructor Alone or an Instructor and a Student”. *CBE Life Sci Educ*.17(2):ar31. doi:10.1187/cbe.17-12-0288
- [24] Kaltura, Powering Any Video Experience, <https://corp.kaltura.com/>. Accessed 2 Mar. 2021.

AUTHORS

Dr. Nguyen is an associate professor in the department of Computer Science and Information Technology at Clayton State University, Georgia USA. Dr. Nguyen has published numerous research articles in the field of Design and Human-Computer interaction. Dr. Nguyen's areas of expertise are in Algorithms, Hi-performance, and Large-scale Computing, Artificial Intelligent, and Full-Stack Software Development. Dr. Nguyen was a software developer prior to joining the academia. His current interests are in the fields of Bioinformatics, Cloud and Mobile Computing, Robotics, and Machine Learning. He published a book on Large-scale Computing in Bioinformatics titled "Biological Multiple Sequence Alignment: Scoring Functions, Algorithms, and Evaluations", John Wiley & Sons, Inc (2016) ISBN-13: 978-1118229040.



Dr. Rahman is a professor in the department of Computer Science and Information Technology at Clayton State University, Georgia USA. Dr. Rahman earned his Ph.D. in computer science from Illinois Tech in the year 2000. Prior to academic appointments he worked as a software developer for the securities industry. Dr. Rahman has been teaching for over twenty years and many of these years he has been trained and an expert in online and remote course development and delivery. His research interests include natural language processing and computer science education. Dr. Rahman has been active in research and publication. He published numerous articles in computational linguistics, user interface design, online learning, and computer science education.

