ConceptScape: Collaborative Concept Mapping for Video Learning

Ching Liu
National Tsing Hua University
Taiwan
chingliu@gapp.nthu.edu.tw

Juho Kim
KAIST
Republic of Korea
juhokim@kaist.ac.kr

Hao-Chuan Wang
National Tsing Hua University, Taiwan
University of California, Davis
hciwang@ucdavis.edu

ABSTRACT
While video has become a widely adopted medium for online learning, existing video players provide limited support for navigation and learning. It is difficult to locate parts of the video that are linked to specific concepts. Also, most video players afford passive watching, thus making it difficult for learners with limited metacognitive skills to deeply engage with the content and reflect on their understanding. To support concept-driven navigation and comprehension of lecture videos, we present ConceptScape, a system that generates and presents a concept map for lecture videos. ConceptScape engages crowd workers to collaboratively generate a concept map by prompting them to externalize reflections on the video. We present two studies to show that (1) interactive concept maps can be useful tools for concept-based video navigation and comprehension, and (2) with ConceptScape, novice crowd workers can collaboratively generate complex concept maps that match the quality of those by experts.

Author Keywords
Crowdsourcing; video learning; online education; concept map.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Videos has become a widely adopted medium for online learning that enables professional instructors or amateur content creators to transfer their knowledge at a large scale. However, video-based online learning environments are still far from an effective place to facilitate meaningful learning. For example, while linking comparative or similar concepts is critical to learning, without close mentoring and support from instructors, novice learners may face difficulty in associating concepts scattered in videos on their own [2]. Also, while it may be enjoyable to watch lecture videos that are carefully crafted and produced, learners may also lose the chance to reflect on what they learned and identify their knowledge gap if relying solely on the video content [10].

Usability problems in video learning environments also prevent learners from effectively organizing their knowledge or retrieving information based on their needs [5, 7]. Linear representation of video is one of the biggest problems that limits learners from exploring learning materials effectively. For example, learners cannot easily navigate a video using concepts of interest to them. Previous research has introduced strategies to improve interactivity of the video player (e.g., [13, 17, 19, 30]), but there's no explicit support for learners to see what concepts are introduced in the video content and how concepts are associated with one another. Visualization is a promising approach to support knowledge exploration [26], and has been applied as a navigation tool for other forms of learning materials [9, 24]. To help online learners escape from the linear mode of video navigation and obtain a concept-oriented view of video learning content, we investigate using concept maps as the visual representation of concepts introduced in lecture videos.

Concept maps as a graph-based visual representation have been used in education for supporting concept communication [20]. Concept maps afford a medium to visually encode concepts on a specific topic possessed by a learner, an expert, or a group of people, and enable peer collaboration [22]. Concept maps may complement lecture videos by offering an abstract, concept-oriented view that's conducive to non-linear navigation and learning of video content.

To better understand if and how concept maps can benefit learners in the context of video learning, we first designed ConceptScape, a prototype interface that integrates a web-based video player with an editable and interactive concept map (Figure 1). The interactive concept map was prototyped to help learners see the conceptual space of the video and navigate the video content through concepts. As we will show in this paper, learners using the prototype reported that concept maps help them comprehend the video on-time and promote reflection afterward, and they may also effectively leverage the maps to retrieve video content based on their needs.

CHI 2018, April 21–26, 2018, Montreal, QC, Canada

ISBN 978-1-4503-5620-6/18/04. . . $15.00

DOI: https://doi.org/10.1145/3173574.3173961

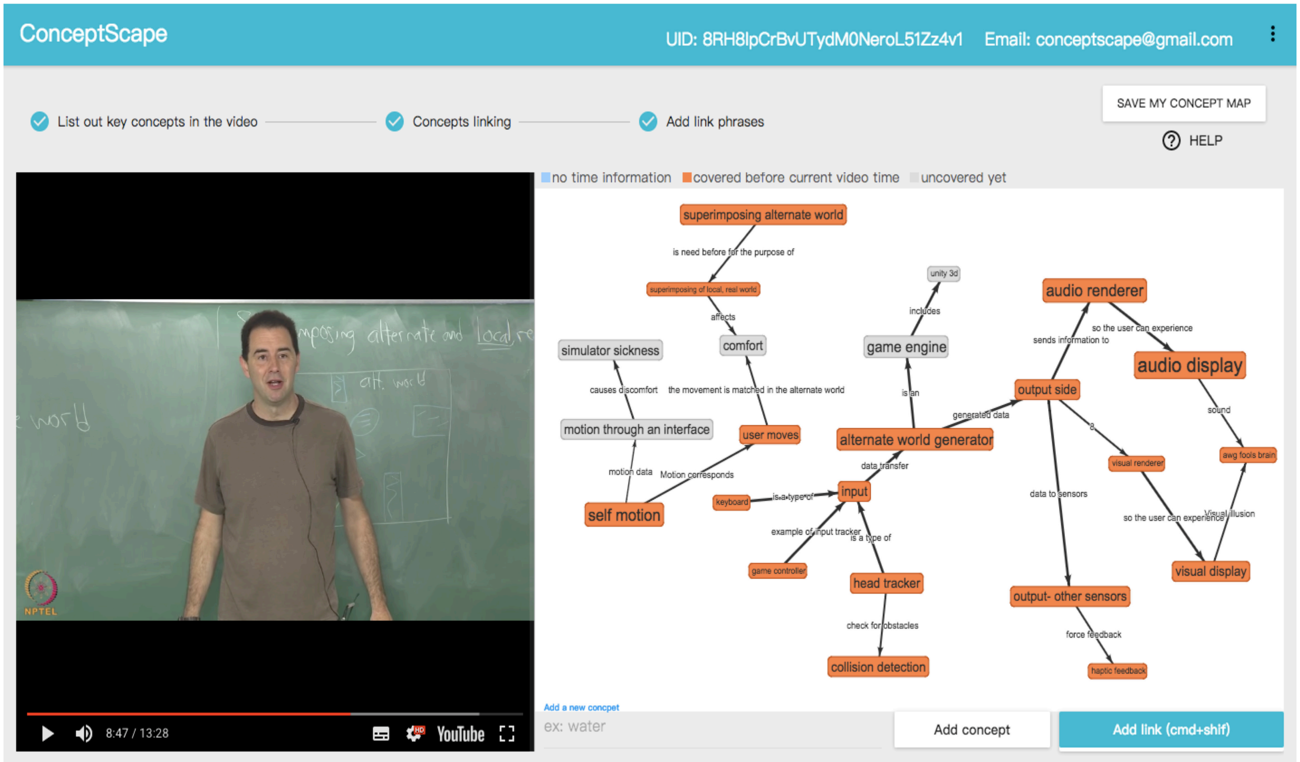

Figure 1. ConceptScape integrates an interactive concept map into a video player. Each concept in the concept map has a time anchor linking to a specific time point in the video, which intends to capture the moment a concept is introduced or explained in the video. The time information is used to visualize lecture progress (by color) and navigate the video (by double-clicking the concept). The size and relative proportion of the video player and the concept map is adjustable, and the space is evenly distributed by default.

However, concept map generation is a complex task, which involves (1) extracting key concepts presented in the lecture, either from audio or visual cues, with their time information, (2) identifying the relations between the concepts, and (3) annotating the relations between concepts with descriptive labels. Limited expert availability presents opportunities and challenges to crowdsourcing. While crowdsourcing this complex task has the potential value to provide support for video-based learning at scale, ensuring the quality of crowdsourced concept maps becomes a challenge.

To scale up concept map support for a large collection of online lecture videos, ConceptScape's crowdsourcing workflow can generate interactive concept maps for any online videos. The crowdsorucing workflow consists of three main stages, following the three main cognitive activities in concept map construction: list concepts, link concepts, and explain relationships. Stages were further divided into steps with different instructions in order to guide workers to focus on different activities in the concept mapping process. A worker may contribute to concept map construction by individually performing one specific concept-mapping activity prompted by the system while watching a video. Our system automatically aggregates the results from the current and previous steps, and advances to the next step or stage.

To evaluate the approach, we recruited participants from Amazon's Mechanical Turk to generate concept maps for three video lectures. Participants were asked to choose a video based on their interest and contribute to one task in the concept map generation workflow. We compared our results to concept maps generated by a learner and by an expert. Results show that collaboratively generated concept maps had comparable quality with respect to expert-generated ones, and better quality than what a single novice learner generated.

The primary contributions of this paper are:

- ConceptScape, a system that integrates an interactive concept map with the video player to support online video learning.
- A crowdsourcing workflow for generating concept maps that can capture workers' reflections on video lectures.
- Experimental results showing that the workflow can generate concept maps that capture a common understanding of video learners with quality comparable with expert-generated versions.

RELATED WORK

As this work introduces a novel combination of video learning and concept maps, we review related work on improving the video learning experience and applying concept maps in online learning.

Research in Video Learning

As a wide variety of video-based learning platforms emerge, learners can access videos that span different subjects and serve different purposes. However, online learners can rarely receive personalized help from instructors, which makes it hard for learners without strong metacognitive skills to construct knowledge meaningfully in online learning [6, 21].

On the other hand, usability problems in video learning environments hinder learners to control their materials [5]. Current research in this area suggests new ways to support learning from instructional videos. First, research attempts to improve video navigation function, which can be within one video (e.g., leverage interaction data for video navigation [13], novice navigation method for blackboard style video [19]) or across videos (e.g., using concept maps to support exploring across learning materials [23]). Other research focuses on scaffolding learners' learning from the content (e.g., crowdsourcing sub-goals labels for how-to video [14]), or improve engagement through exercises (e.g., [15, 12]). Extending the previous research, we focus on integrating comprehension support and concept-based learner-video interaction into video learning environments.

Concept Maps in Online Learning

A rich body of previous research has shown the value of providing knowledge representation (e.g., concept maps, mind maps, or knowledge graphs) in online education. Wang et al. investigate knowledge visualization to support resource-abundant and self-regulated text-based online learning [26]. Teo et al. propose a knowledge-driven model to personalize e-learning [25]. Schwab et al. explore hierarchical concept maps to support dynamic non-linear learning planning for modularized short videos [23]. These approaches mainly utilize the non-linear property of map- or graph-based knowledge representation to support organization of learning materials across different resources, which is critical for self-regulated learning [31]. In this research, we expect such positive effects of concept maps can be applied to support organizing and indexing concepts within a video.

CONCEPTSCAPE: CONCEPT MAP INTEGRATED INTO VIDEO PLAYER

To understand if and how concept maps can benefit learners in the context of video learning, we first designed ConceptScape, a prototype interface that integrates a web-based video player with an editable and interactive concept map (Figure 1). Each concept in the concept map has a time anchor linking to a specific time point in the video, which intends to capture the moment a concept is introduced or explained in the video. Although a concept may appear in different parts of the video (e.g., showing examples of a concept applied in practice, using a concept to introduce another concept), ConceptScape captures a moment where the instructor provides an explanation of the concept. Learners can navigate to that moment by double clicking on the concept. ConceptScape also visualizes in-video progress by changing the color of concept nodes that have already been covered (gray to orange), which is analogous to the standard linear video timeline that visualizes the current in-video time. By visually distinguishing the covered and upcoming concepts, we hope to encourage learners to reflect on the concepts and their relationships.

To support learners in organizing their understanding and customizing their learning material, we allow learners to edit the concept map. They can add, update, and delete concepts, links, or link phrases. In terms of usability, several considerations are taken into account to support learners watching a video and interacting with a concept map simultaneously: zooming in/out and panning are supported for easier navigation, and keyboard shortcuts are added for efficient concept map editing.

Pilot Study: Learners' Use of Concept Maps in Video Learning

To evaluate the effects of using ConceptScape for video learning, we conducted a pilot study with online video learners.

Study Design

We selected two videos in different topics (C1: Introduction to Software of Virtual Reality[1], C2: What is Gamification?[2]) that have similar lengths (C1: 13:28 and C2: 11:51). Participants rated difficulty of each video after watching them through 10-point Likert scale (1:very easy, 10:very difficult), and results show similar perceived difficulty between the two videos (C1: M=4.3, SD=2.5 and C2: M=5.3, SD=1.6 ($t(18)=1.073$, $p=0.298$)). For each video, we provide a handcrafted concept map generated by the first author. The concept maps were further modified by removing some elements because we hoped to understand whether a concept map is helpful even if it's incomplete. The concept maps had different levels of complexity (C1: 9 concepts, 8 links, no link phrase; C2: 24 concepts, 20 links, 12 link phrases), reflecting the complexity of lecture content.

Participants

We recruited 20 participants [P1-P20] (10 male and 10 female) through online social media posting. Most of participants were college students. They received $3.3 for up to 30 minutes of participation.

Task and Procedure

The study was conducted online. Participants were required to visit our website and watch a video lecture with a preconstructed concept map. Once a participant visited our website, they were randomly assigned to watch one of the two videos.

Participants were asked to improve the concept map according to what they learned, but there were no further constraints on the improvement task. This was to ensure that every participant at least sees the concept map to some degree and has freedom in their use of the concept map. After they finished watching the video and improving the concept map, they were asked to answer a questionnaire.

The questionnaire included questions to understand their self-evaluated background knowledge about the lecture topic, the perceived difficulty of the video, their understanding about the lecture, and the level of engagement. Next, there were three open-ended questions, asking participants about their

[1] `https://www.youtube.com/watch?v=1LpHDOWMAdA`
[2] `https://www.youtube.com/watch?v=BqyvUvxOx0M`

learning experience with the concept map before, during, and after watching a video, followed by questions probing their experience on concept map editing.

Results

Participants reported limited prior knowledge on the assigned topics (10-scale Likert question: 1: never heard about it, 10: understand very well), with C1: M=4.9 (SD=3.07) and C2: M=3.0 (SD=2.75). Their self-evaluated engagement while watching the video was high in both cases (10-scale Likert question: 1: not engaging, 10: very engaging), with C1: M=8.6 (SD=0.84) and C2: M=8.0 (SD=1.82). Participants' self-evaluated understanding after watching the video were moderately high (10-scale Likert question: 1: don't understand anything, 10: understand really well), with C1: M=6.8 (SD=2.4) and C2: M=6.5 (SD=1.84).

Overall, learners reported seeing concept maps before watching the video didn't help their understanding that much. When learners didn't have enough domain knowledge [P4], or encountered unfamiliar concepts [P3, P6] and connections [P5], they considered it difficult to comprehend the concept map. Furthermore, since most learners were not familiar with using concept maps, it was hard for them to interpret the knowledge behind a structural representation [P10, P20]. On the other hand, some learners found the concept map helpful before watching the video because it provided a useful summary [P1, P16] and an overview, helping them identify important topics beforehand [P2, P8, P15, P17].

Most learners considered watching video along with a concept map helpful. Some learners thought the concept map was a cognitive road map that helped them be aware of the ongoing section through distinguishing colors and see connections to other sections [P7, P11, P17]. Concept maps helped learners follow along instructions in the video [P8] and clarify their knowledge [P5, P10, P7]. Some remarked that they used the concept map to organize their notes [P6, P20].

Reviewing a concept map after watching the video was considered helpful because it reinforces learners' understanding of the lecture by promoting comprehensive recall [P6, P7, P10, P19], reflection on [P1, P10] and summarization of [P7, P8, P18] the content. While performing these cognitive activities with a concept map, learners said they were able to quickly refer back to the video section for clarifying and relearning unclear concepts [P2, P5, P10].

Finally, editing an existing concept map while watching the video was commonly considered valuable. The activity reinforced learners' memory [P1, P6] and understanding [P9]. Most of all, learners expressed a positive learning experience when they found discordance between the existing concept map and their mental model, such as "I tried to change the concept map into a way that I understand" [P6], "I rethink what I learned and digest those content into my knowledge" [P7], "I feel it can promote my ability to organize and recap my thought" [P15]. Learners also mentioned making improvements on the concept map gives them a sense of accomplishment [P13, P20]. These feedback echoes findings from previous research in concept mapping [20]; that is, fostering meaningful learning where learners assimilate new knowledge with their existing knowledge.

In our preliminary qualitative results, learners reported that ConceptScape could effectively support their understanding of the lecture and reflection on their knowledge. The interface also provided them a shortcut to refer back to the specific section when they found anything unclear. However, it requires future work to understand how large of a learning gain and what other types of scaffolding interactive concept map may afford.

Participants also expressed encountering difficulties in editing a concept map, such as distraction from video, confusion in using the interface, and ambiguity in how to improve the concept map (e.g., "am I adding too many details?"). In the second study, we improved the interface of our system for crowdsourcing concept map generation for online lecture videos at scale.

CROWDSOURCING CONCEPT MAP GENERATION: MOTIVATION AND DESIGN GOALS

Generating a concept map that captures major concepts from a lecture video is an interdependent and complex process. Automated concept map generation methods mainly rely on Natural Language Processing (NLP) techniques [32]. To generate concept maps for video contents, an inherent challenge stems from dealing with audio and visual tracks at the same time. Automated video concept map generation may be achieved by converting auditory and visual information to text through speech recognition and computer vision [32]. But these techniques cannot be easily generalized to lecture videos that may contain diverse visual representations (e.g., animation, hand writing) or complex audio sources (e.g., discussion between the instructor and students). Experts' manual generation can yield quality concept maps, but it suffers from limited scalability.

We propose collaborative concept map generation by learners as an alternative to the aforementioned approaches. Asking learners to generate a concept map may also potentially provide a learning opportunity by encouraging them to verbalize and summarize concepts. For quality control, a well-designed collaboration process could arguably filter out incorrect input from individual novice learners. In addition, we can further compare concept map components from multiple learners and collect agreement information as metadata, for example, showing how many people added the same link in their concept maps.

Although learners are a plausible population to construct concept map, it is difficult for individual novices without sufficient domain knowledge and concept mapping experience to generate concept maps by themselves [4]. Research also shows that collaborative concept map generation can produce higher quality result [16], since the process of seeing how others interpret the content through an externalized knowledge representation can promote collaborative learning. In traditional classroom settings, collaborative concept mapping in small groups is a well-explored pedagogical strategy [20]. However, to our knowledge, there has been no attempt to organize large-scale online learners into a collaborative concept map construction process, let alone collaborative concept mapping for video content representation.

COLLABORATIVE CONCEPT MAP GENERATION: CROWDSOURCING WORKFLOW AND INTERFACE

We designed a crowdsourcing workflow for concept map generation (Figure 2). The design goal of our crowdsourcing concept map interface is to support workers to generate and edit concept map components while giving them a chance to reflect and organize their knowledge and minimizing distraction from video watching. We took an iterative design process, by running informal pilot studies to get comments from workers. Through iterations, we made designs decisions to guide the workers' concept mapping process, such as: (1) provide a workspace for moving around and grouping concepts (e.g., canvas style workspace can promote workers to find out implicit connections between concepts), (2) support easy recall and playback of concepts (e.g., ordering the concepts by their time of appearance in video and helping workers refer to video when they construct a concept map).

ConceptScape's crowdsourcing workflow design extends that of multi-stage workflows (e.g., Soylent [3], ToolScape [11], Crowdy [27]), which are useful in dividing a large, complex task into smaller units for crowd workers. Following previous research in crowdsourcing, we apply similar design of multi-stage workflow for concept map generation. The main challenge we are tackling for generating a concept map, which can capture workers' common understanding on the video lecture, is to balance the trade-off between giving space for individual reflections and setting constraints to foster consensus.

Our crowdsourcing workflow for concept map generation has three stages: concept and timestamp generation, concept linking, and link labeling. Each stage is designed to yield different types of output, and within a stage multiple steps are added for quality control. Each stage has a unique interface dedicated to handling certain components and each step has a specific instruction for the task. When a worker accesses the workflow, she is assigned a step within a stage, and is given a partial concept map aggregated from the available results so far. Our first design insight is to enable workers to contribute in parallel (higher efficiency) while maintaining sequential step transitions (better quality control). That is, workers within one step cannot see each other's work and make parallel progress, but upon collecting enough data, the system aggregates the results and advances to a next step for the next group of workers to work on. Second, we afford workers more natural tasks of concept mapping by giving them flexibility to perform multiple tasks within a stage (e.g., link generation) instead of restricting to a step (e.g., pruning link only). The flexibility of task design helps us gather extra contributions, which can improve result quality. Now we describe the workflow in detail.

Stage 1: Concept and Timestamp Generation

In this stage, the workflow aims to produce a collection of concepts within a lecture with their timestamps. We improve the quality of concepts by dividing this stage into three steps: Find-Prune-Adjust, to collect abundant concepts while removing duplicates and to obtain correct timestamps pointing to the sections the instructor is explaining about. ConceptScape gives workers a tool (Figure 3) that allows them to write down key concepts while watching a video. Right beneath the input area, we show workers an explanation on 'what makes a good concept' to promote them to contribute short and critical concepts. When a worker adds a new concept by clicking on the 'Add' button or pressing the enter key, the system gets the current video playback time minus 3 seconds as a timestamp attached to the new concept. The 3-second time buffer is assigned to compensate the typing period before a concept is added. The timestamp supports navigating the video by clicking on the concept, which also allows simple adjustment to re-anchor the timestamp to another video time. Concepts are ordered by their timestamps based on the workers' needs we identified in our pilot study.

Find Step: *Please write down key concepts that the instructor is teaching.*

In this step, workers get an empty concept list and are asked to fill up their own. To aggregate the results, we use an automatic document clustering API provided by MeaningCloud [1] to group the concepts by their meaning and use DBSCAN (Density-Based Spatial Clustering of Applications with Noise) to identify the highest density of time clusters for generating timestamps. The server gathers all concepts from workers and sends a cluster request to MeaningCloud [1], which returns a result containing multiple groups of concepts. The system then removes groups that have less than three concepts (agreement threshold). For each group, the system records how many items are included as a measure of agreement. Then, it goes through two sorting processes to identify a representative term: (1) sort the concepts by their length (shorter concepts get higher priority), (2) count the number of identical concepts and sort the concepts by the count (commonly written concepts get higher priority). In short, a concept is likely to be identified as a representative if it is short and/or contributed by more workers. To assign timestamps to concepts, we follow the same method used for deciding video event timing [14], which clusters timestamp candidates into groups and identifies the one with highest density. We pick the earliest timestamp as the representative concept, in order to capture the beginning of an explanation for that concept as much as possible. Finally, a list of concepts containing their representative terms (labels), timestamps, and agreement frequency is saved to our database.

Prune & Adjust Step: *Please delete duplicate or unnecessary concepts. & Please focus on adjusting the 'time' of each concept to when the instructor is explaining it.*

When a worker comes to a course, they get an initial concept list with their timestamps produced from Find Step, and they can also use the concepts to navigate the video. The worker is asked to prune concepts or adjust timestamps depending on which step it is, but they are not limited to pruning or adjusting since our tool affords them to do all the tasks in this stage (add, edit, and delete both concepts and timestamps). Giving

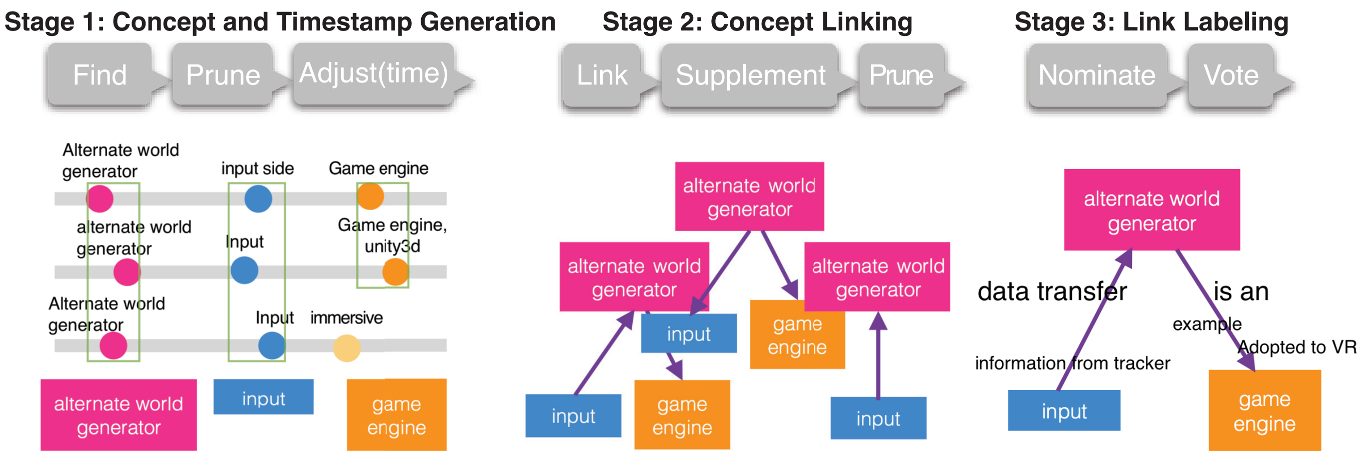

Figure 2. ConceptScape's crowdsourcing workflow includes three stages with eight detailed steps. Workers perform micro concept mapping tasks parallel in each step and our system automatically aggregates collected contributions within a step and propagate the result to the next step. Steps are divided for quality control, guiding workers to focus on certain task but not restricted to. We allow workers to work on different tasks within a stage to make the task more natural for learning and also help us gather extra contributions.

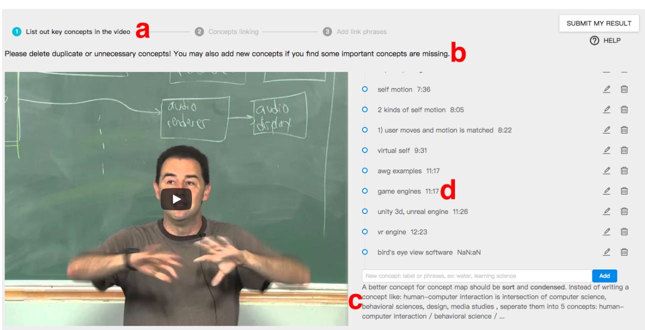

Figure 3. The ConceptScape crowdsourcing interface shows workers (a) stage progress and (b) a specific instruction guiding them what they should do in the current step. In Stage 1, workers collaboratively generate a set of concepts. It (c) presents an explanation about what makes a good concept and (d) automatically attaches timestamps to newly added concepts. They can click on the concepts to navigate the video.

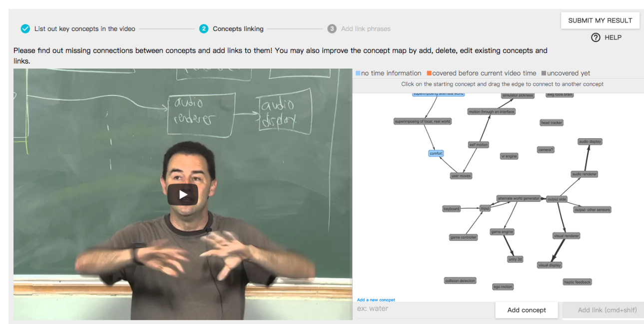

Figure 4. In Stage 2, workers drag concepts and draw or delete links to construct, supplement, or prune connections.

workers this flexibility is based on two reasons: (1) improving overall result quality (2) affording a more natural task to fit to their learning process. Since each step is highly dependent to previous steps, the quality yielded from previous steps can influence next-step workers' result and also working experience. For example, if the concept list from Find Step contains no redundant concepts after machine aggregation, then limited workers in Prune step to prune concept may force them to do random deletion that can hurt our quality. On the other hand, limited workers to only pruning or only adjusting timestamp also forbids them to use the tool to support their learning (writing down important concepts while watching video). The same aggregation method used in Find Step is used in these two steps.

Stage 2: Concept Linking

Based on the extracted concepts from Stage 1, this stage collects connections between concepts. A connection is represented by an arrow linking two concepts, and we use the width of an arrow to show the popularity of that connection. This stage contains three steps: Link-Supplement-Prune. Workers use a concept mapping workspace to draw links and arrange structure. Workers also see colors on concepts that reflect playback status and navigate the video with concepts. The crowdsourcing interface is shown in Figure 4.

Link Step: *Please link the concepts to visualize the structure and idea of this lecture. You may also add, delete, or edit existing concepts.*

When a worker enters this step, they see a set of concepts (an aggregated result from Stage 1) linearly ordered by their timestamps on the concept mapping workspace (Figure 4, right). Workers need to move concepts and make links. We also allow workers to change the concepts if they want. To aggregate the result, we first use the same method in Stage 1 to cluster concepts into semantic groups. After getting groups of concepts with their representative labels, all concept-to-concept links are counted as group-to-group links. For example, if 'input sensor' belongs to the 'input' (representative label) group and 'output device' belongs to the 'output' group, then a link from 'input sensor' to 'output device' is treated the same as a link from 'input' to 'output'. We aggregate those processed links and count their agreement numbers. Links without any agreement (links only identified by one worker) are removed, while other links with concepts are saved in our database.

Supplement Step: *Please find out any missing connections between concepts and add links to them. You may also improve the concept map by adding, deleting, or editing existing concepts and links.*

In this step, workers sees an initial concept map from Link Step. All links in the initial concept map have at least two worker's agreement, and the width of a link gets thicker when it has more agreement. Building upon on a partially constructed graph, we expect workers in this step to get more consensus than workers in Link Step. We use both data from Supplement Step and Link Step as raw data for aggregation. The aggregation method greedily collects all plausible connections, so we use more data while increasing minimum agreement (=3 in our experiment, 2 in Link Step) as threshold. For example, if a link is only contributed by one worker in Link Step but gets two more workers making the same link, it is counted as a valid link. On the other hand, if a link already exists in the initial concept map, then it is almost impossible to be deleted in this step. We use the similar method to Link Step for link aggregation, but instead of using text clustering for pre-processing the concepts, we simply take original terms to represent a link; that is, a more restrictive concept aggregation method is adopted here since we intend to converge the concept. A new concept can be added only if it is connected to another concept by more than three people using the same terms. While allowing for flexible work between steps, we enforce increasing quality control in later steps. This design choice is made to ensure that concepts get more than a single chance to be included in the final concept map while gradually having to meet a higher standard.

Prune Step: *Please delete unnecessary links to make the concept map clearer. You may also improve the concept map by adding, deleting, or editing existing concepts and links.*

While Link Step and Supplement Step aim to discover possible links, this step finalizes the links by pruning. Workers are asked to prune the links but not restricted to only pruning. We use raw data from Supplement Step and Prune Step and apply the same aggregation method in this stage, but agreement threshold is set to be the same number of workers (=10 people in our experiment) in this step. The agreement threshold means that we require same number of people who delete a link in Prune Step to be the same as the number of people who has that link in Supplement Step. For example, if a link exists in 5 concept maps of workers in previous step, then we require at least 5 removal here to delete the link. Notice that a link from Supplement Step could be new added one, or remaining one from Link Step. A link generated from Link Step usually has higher agreement than new added one from Supplement Step, because we expect a concept map aggregated from Supplement Step contains most of the links from Link Step. Therefore, it requires more deletion from Prune Step to delete a link that comes from Link Step. This design is to ensure pruning quality by balancing the effort of adding a new link and deleting an existing one.

Stage 3: Link Labeling

The goal of this stage is for workers to label the links (i.e., link phrases) generated in Stage 2. Workers do this either by verbalizing the relationship in their words or voting for link phrases from others. Naturally, this stage is divided into two steps: Nominate and Vote.

Nominate Step: *Please add labels(link phrases) to the links to verbalize the relationship between concepts.*

In this step, workers verbalize the relationship between concepts and label the links. Updating the links themselves is also allowed, since workers might find it necessary as they review the concept map. We use the same method in Supplement Step of Stage 2 to aggregate link changes, and additionally collect possible link phrases for each link. After collecting nominated link phrases, our system aggregates link phrases by counting duplicate terms as agreement counts and then keeps unique terms. The link phrases are further sorted by their agreement counts. After aggregation process, the final nominated link phrases for each link will be a list of unique terms sorted by their agreement of nominators.

Vote Step: *Previous workers have added some link phrases. Click on the links and choose a link phrase that best verbalizes the relationship.*

Workers see a concept map with nodes and links when they come to this step. The link phrases added by workers from previous step are embedded into links. When workers click on a link a list of nominated link phrases will show up, so they can pick the best link phrase for each link. Every worker has to review all the links and they will see a concept map with link phrases on each link when they finish the task. Our system later selects the link phrase with most votes as a representative link phrase. The system finally generates a complete concept map after this step.

EVALUATION

To evaluate if the ConceptScape workflow can be used by remote, independent users online to generate a high-quality concept map, we conducted an online experiment. We focused on evaluating the overall quality of concept maps generated by our workflow. We further investigate the quantity of individual contributions on the tasks to understand how workers contribute to the tasks when given a flexible tool.

Participants and Materials

We deployed ConceptScape online and embedded detailed instructions telling users how to use our system and an overview of the crowdsourcing process. We recruited crowd workers from Mechanical Turk to fully construct concept maps for three lecture videos. The videos varied in topics and representation styles (see Table 1). Overall, 180 HITs were published, and we rewarded $4 for each HIT. Each video required 60 HITs, and we heuristically allocated the HITs to the three stages : Stage 1 (Find-Prune-Adjust): 10-6-6, Stage 2 (Link-Supplement-Prune): 8-8-10, and Stage 3 (Nominate-Vote): 6-6.

Each worker could contribute to multiple videos but could not contribute to the same video multiple times. For the

180 HITs, we had 123 unique contributors. Workers' demographic information and their prior level of knowledge on the topic, and familiarity with concept mapping are summarized in Table 2.

Video (abbreviation)	Presentation Style	Length
Hello World Machine Learning [a] (ML)	Talking heads with animated slides	6:52
Introduction to Software of Virtual Reality [b] (VR)	Classroom lecture	13:28
Why is being scared so fun? [c] (TED)	Ted-Ed animation	4:28

Table 1. To evaluate our workflow, we created concept maps for three videos using the workflow.

[a] `https://www.youtube.com/watch?v=cKxRvEZd3Mw`
[b] `https://www.youtube.com/watch?v=1LpHDOWMAdA`
[c] `https://www.youtube.com/watch?v=oetVvR5RQUs`

Video	Gender (Female, Male)	Education level ($\leq$High school, $\leq$B.S., $>$B.S.)	Self-reported prior knowledge level: M (SD)	Concept mapping familiarity: M (SD)
ML	28, 32	2, 51, 7	2.84 (1.66)	3.09 (1.75)
VR	30, 30	6, 46, 8	2.89 (1.74)	2.82 (1.81)
TED	31, 29	5, 45, 10	1.78 (0.93)	2.22 (1.31)

Table 2. Participants' demographics, prior level of knowledge on the topic (7-point Likert scale: 1: without any background knowledge, knowing about this topic very well), familiarity with concept mapping (7-point Likert scale: 1: without any experience, 7: high familiarity).

Analysis

Quality evaluation: To evaluate the quality of system-generated concept maps, we compared our results to concept maps generated by domain experts and individual novices, respectively. Three expert-generated concept maps were produced by the first author (TED) and another domain expert (VR, ML) with prior concept mapping experience. Three novice-generated concept maps were produced by three college students. They were given an example concept map with instructions. Both experts and novices used the ConceptScape interface (Figure 4) to construct a complete concept map by themselves.

Finally, we invited another group of two domain experts for each video to evaluate the quality of concept maps generated from the three conditions: from an expert, from a novice, and from ConceptScape. In a blind condition, the evaluators watched the video and scored concept maps independently according to a provided scoring rubric, which included the following components:

- **Holistic evaluation**: Evaluators rate in 1-10 to indicate the overall quality of a concept map (adopted from [18]).
- **Component evaluation**: Evaluators score three components separately, namely concepts (if valid, give 1 point), links (if valid, give 1 point), and link phrases (if valid, give 1 point), and sum up three component scores to a total component score (adopted from [16]).

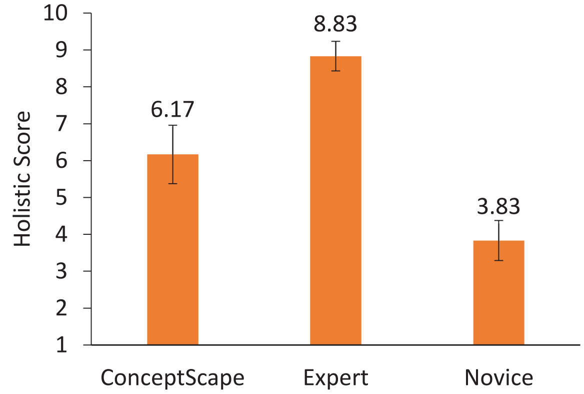

Figure 5. Mean of the holistic scores of concept maps produced by ConceptScape, expert, and novice. The result shows that our crowdsourcing concept maps have quality that holistically rated a little bit lower than experts' generated ones but much higher than novices' individually generated results.

The interrater reliability between the two evaluators is calculated with Pearson's correlation between their scores (holistic score and component score) on three concept maps. The interrater reliability correlations in ML, VR, TED were 0.91, 0.96, 0.98, respectively, indicating high reliability.

Quantity of contribution: Individual contributions include main contribution (main assigned task) and extra contribution (additional work). Analyzing workers' individual contributions is important because we did not pay bonus for extra contributions and it is indicated in our instruction; however, their extra contribution can indead improve the result due to our workflow design and aggregation method. The quantity of contribution is calculated as the sum of newly added or edited components in a concept map. For example, for a worker is assigned a pruning task in Stage 2, if he deletes 4 links, adds 6 concepts, and edits 8 links, then he has made 4 main contributions and 14 extra contributions (which is 6+8).

RESULTS

Concept Map Quality

We demonstrate one of our crowdsourcing concept maps in Figure 1 (for VR video). Across all videos the mean holistic scores (Figure 5) from the three conditions were: 8.83 (Expert), 6.17 (ConceptScape), and 3.83 (Novice), and the mean total component scores (Figure 6) were 58.5 (Expert), 47.17 (ConceptScape), and 14 (Novice). One-way ANOVAs showed a significant main effect of condition (Expert, ConceptScape, Novice) on holistic score ($F(2,6)=9.8$, $p=.013 < .05$) and a significant main effect of condition on total component score ($F(2,6)=21.92$, $p<0.01$). Table 3 shows average holistic scores from the two raters on each concept map. Table 4 shows average three component scores (concept, link, link phrases) of each concept map.

We further compare the ConceptScape group to Novice group and Expert group. **ConceptScape vs. Novice:** A paired t-test

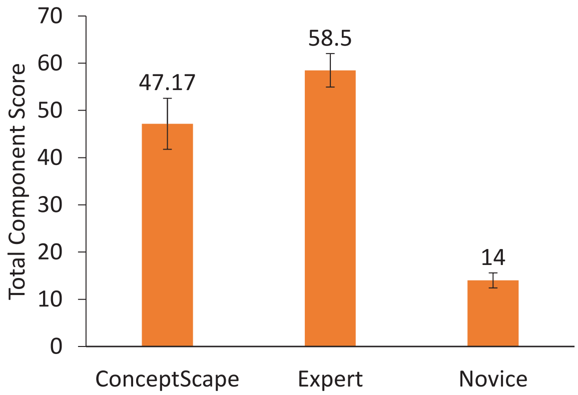

Figure 6. Mean of the total component scores of concept maps produced by ConceptScape, expert, and novice. The result shows that our crowdsourcing concept maps contain more valid and sufficient components than novice generated ones, and their component-level quality are almost as high as expert generated concept maps.

revealed significant difference in the total component score between groups (t(4)=4.67, $p<0.01$), but no significant difference in the holistic score (t(4)=1.78, p=.15). **ConceptScape vs. Expert:** A paired t-test showed no significance in both the holistic score (t(4)=2.26, p=.09) and the total component score (t(4)=1.39, p=.24).

To evaluate per-stage performance of the workflow, we investigate the component score of concepts, links, and link phrases (Figure 7). For all components, ConceptScape produced significantly higher quality components than the Novice group (concepts: t(4)=6.07, $p<0.01$; links: t(4)=3.89, p=0.02; link phrases: t(4)=3.73, p=0.02). On the other hand, there is no significant difference between ConceptScape and the Expert group in terms of the component scores (concepts: t(4)=0.32, p=.77; links: t(4)=1.42, p=.23; link phrases: t(4)=1.74, p=0.16). This suggests that individual concept map components generated by ConceptScape have comparable quality to expert-generated ones.

In summary, our results show that ConceptScape generated concept maps with comparable quality to expert-generated concept maps, in terms of both the holistic evaluation and the component level evaluation.

	ConceptScape	Expert	Novice
ML	4	9	2.5
VR	7	8	4
TED	7.5	9.5	4

Table 3. Summary of holistic scores for all concept maps.

	ConceptScape			Expert			Novice		
	C	L	P	C	L	P	C	L	P
ML	16.5	11	8.5	18	15.5	15.5	6.5	5	0
VR	21.5	19.5	18	22	21.5	20.5	9	6.5	3.5
TED	17.5	15	14	17.5	22.5	22.5	5	3	3.5

Table 4. Summary of component scores (C), links (L), and link phrases (P) for all concept maps.

Individual Contribution

We found extra contributions in all steps (Step 1 in Stage 1 is excluded from the analysis since no extra contribution was available to workers), indicating that workers indeed contributed to the concept map much more than they were required to. A summary of total main and extra contributions is reported in Table 5. Since the interface affords workers to organize their knowledge flexibly, and concept map generation steps are intuitive, their extra voluntary contributions suggest that they were motivated to participate in the concept mapping activity while watching a lecture video. While we conducted the study on a crowdsourcing platform with monetary reward, observing spontaneous contributions from our participants implies a potential to ask unpaid online video learners to collaboratively construct a concept map.

Note that we found tasks related to pruning receiving more extra contributions than the main contributions (highlighted rows in Table 5). This may be because the initial concept map (or concepts) didn't require much pruning work, or likely the aggregation algorithm may have already removed a significant portion of noisy data.

Stage	Step	Main task	Extra task
1	Find	308	-
	Prune	44	108
	Adjust	107	111
2	Link	442	126
	Supplement	181	77
	Prune	38	229
3	Nominate	373	79
	Vote	195	75

Table 5. The amount of individual contributions in each step.

DISCUSSION AND LIMITATIONS

We conducted a pilot study to gather preliminary feedback from learners about their experience of video learning with an interactive concept map. While positive feedback demonstrates a potential that ConceptScape may support learning from video, it is still unclear in what degree can a concept map improve content learning. Furthermore, a concept map also may serve as an extra video navigation tool. It is worth investigating the three-way interaction between learners (their intellectual process), video content (original learning material), and the concept map (structured knowledge representation from others) in the future.

We followed the most typical design of concept maps (demonstrated in [20]), including the elements (nodes, links, link phrases) and their forms. However, there may be other interface features that will further improve usability and decrease the burden of interacting with a concept map (e.g., tedious drag-and-drops). An improved design may also decrease the complexity of concept maps to lower learners' distraction. Another extension of this work can go beyond linking concepts within a video, but instead linking concepts across different media and sources of information.

Recent research has introduced approaches to engage learners in producing useful learning materials, known as learnersourcing. Learnersourcing applications have been introduced

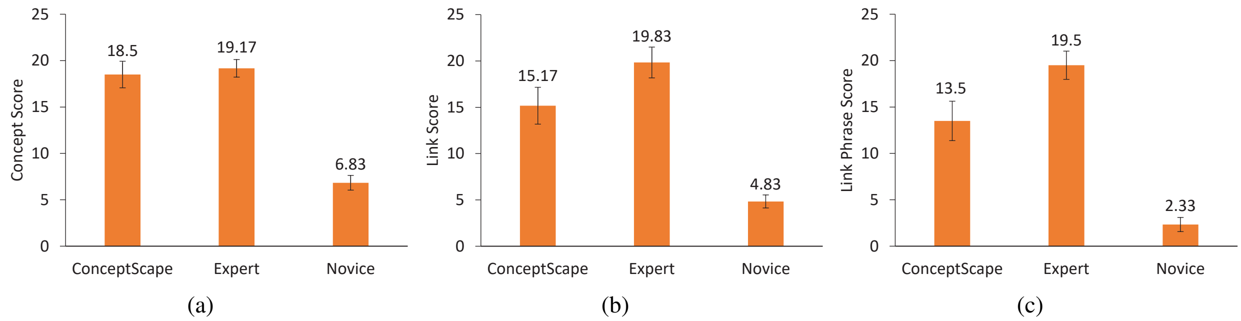

Figure 7. Mean of component score of (a) concepts, (b) links, and (c) link phrases

to extract subgoals for how-to videos [27], to generate text-based explanations for answers to questions [29], to generate personalized hints for problem-solving [8], and to collect video-based explanations to demonstrate how to solve a math problem [28]. These examples demonstrate an opportunity to organize a large number of learners online into collective interactive content generation.

While the current paper involves crowd workers, not regular learners who're motivated and prepared to learn, we found some potential to involve a group of learners into ConceptScape's collaborative concept mapping workflow. First, from the evaluation of the quality of crowdsourcing concept mapping, it's clear that crowd workers without expertise on the subject matter and without much experience in concept mapping could generate content of similar quality as the work of an expert. While an individual novice learner might lack such capability, it is plausible that we can replace crowd workers with a group of learners in the workflow to achieve the same or better quality of concept mapping, not to mention that learners are likely to be more self-motivated and knowledgeable than random crowd workers recruited online. Second, the emergence of spontaneous extra contributions from workers suggests that our task design could be natural to video learners. To involve online learners into collaboraive concept map construction, as a next step we will focus on verifying the learning benefits for "learner workers".

ConceptScape's three-stage crowdsourcing workflow includes a total of eight steps. We derive these steps based on observations obtained through design iterations. The main challenges we are tackling is to elicit more individual reflections, which is beneficial for workers' own learning, while reaching consensus among workers. We promote workers to externalize their own reflections by putting them to work in parallel in each step and provide flexibility in their work. While it is hard to aggregate concept maps constructed in parallel, we divide each stage into multiple steps and show current collective results in a stage. Though our workflow demonstrates its ability to generate quality concept maps, further studies are required to inform decisions for dividing the tasks, which is based on heuristics at this point. It would also be helpful and interesting to investigate how much flexibility is required in learnersourcing concept mapping to achieve the balance between quality content generation and learning through reflection.

In summary, ConceptScape generates concept maps through capturing individuals' reflections on a video when they perform micro concept mapping tasks. Our aggregation method records workers' agreement on concept map components (e.g., concepts, links, and link phrases) and reveals the population of agreement through the size of concepts, the width of links, and the presented order of link phrases (Figure 1 shows a crowdsourced concept map). Visualizing these information with interactive features may help online learners locate important parts in a video, and increase the awareness of the confidence on crowdsourcing the generation of specific elements of a concept map.

CONCLUSION

This paper presents ConceptScape, a system that generates and presents a concept map for lecture videos. We introduce a crowdsourcing workflow to engage workers to collaboratively generate a concept map by prompting them to externalize reflections on the video. We evaluate our crowdsourcing workflow on Mechanical Turk. The result shows that crowd workers collaboratively generated concept maps that match the quality of those generated by experts. In addition, the flexible task design of the workflow promotes workers to contribute more than required, while they generally perceived performing the task to be helpful for learning. We also show that watching video with an interactive concept map can support concept-based video navigation and comprehension.

ACKNOWLEDGMENTS

The authors thank Shun-Huai Yao for his support in collecting and analyzing experimental data, and members of the NTHU CSC Lab and KAIST KIXLAB for their feedback. This research was supported in part by Ministry of Science and Technology of Taiwan (MOST 106-2633-E-002-001, 105-2628-E-007-004-MY2), National Taiwan University (NTU-106R104045), Intel Corporation, and Next-Generation Information Computing Development Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C4A7065960).

REFERENCES

1. MeaningCloud: Text Analyze API. `https://www.meaningcloud.com/developer/apis`. (Accessed Aug. 2017).

2. Mohamed Ally. 2004. Foundations of educational theory for online learning. *Theory and practice of online learning* 2 (2004), 15–44.

3. Michael S Bernstein, Greg Little, Robert C Miller, Björn Hartmann, Mark S Ackerman, David R Karger, David Crowell, and Katrina Panovich. 2015. Soylent: a word processor with a crowd inside. *Commun. ACM* 58, 8 (2015), 85–94.

4. K.E. Chang, Y.T. Sung, and S.F. Chen. 2001. Learning through computer-based concept mapping with scaffolding aid. *Journal of Computer Assisted Learning* 17, 1 (2001), 21–33. DOI:**http://dx.doi.org/10.1111/j.1365-2729.2001.00156.x**

5. Konstantinos Chorianopoulos and Michail N Giannakos. 2013. Usability design for video lectures. In *Proceedings of the 11th european conference on Interactive TV and video*. ACM, 163–164.

6. D Randy Garrison and Martha Cleveland-Innes. 2005. Facilitating cognitive presence in online learning: Interaction is not enough. *The American journal of distance education* 19, 3 (2005), 133–148.

7. Michail N. Giannakos, Konstantinos Chorianopoulos, Marco Ronchetti, Peter Szegedi, and Stephanie D. Teasley. 2013. Analytics on Video-based Learning. In *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK '13)*. ACM, New York, NY, USA, 283–284. DOI: **http://dx.doi.org/10.1145/2460296.2460358**

8. Elena L. Glassman, Aaron Lin, Carrie J. Cai, and Robert C. Miller. 2016. Learnersourcing Personalized Hints. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1626–1636. DOI: **http://dx.doi.org/10.1145/2818048.2820011**

9. Katsuhiro Ikeda, Kozo Sugiyama, Isamu Watanabe, and Kazuo Misue. 2006. Generation of Relevance Maps and Navigation in a Digital Book. In *Proceedings of the 2006 Asia-Pacific Symposium on Information Visualisation - Volume 60 (APVis '06)*. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 49–58. **http://dl.acm.org/citation.cfm?id=1151903.1151911**

10. Heather Kanuka and Terry Anderson. 2007. Online social interchange, discord, and knowledge construction. *International Journal of E-Learning & Distance Education* 13, 1 (2007), 57–74.

11. Juho Kim. 2013. Toolscape: enhancing the learning experience of how-to videos. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2707–2712.

12. Juho Kim, Elena L. Glassman, Andrés Monroy-Hernández, and Meredith Ringel Morris. 2015. RIMES: Embedding Interactive Multimedia Exercises in Lecture Videos. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 1535–1544. DOI:**http://dx.doi.org/10.1145/2702123.2702186**

13. Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014a. Data-driven Interaction Techniques for Improving Navigation of Educational Videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, USA, 563–572. DOI:**http://dx.doi.org/10.1145/2642918.2647389**

14. Juho Kim, Phu Tran Nguyen, Sarah Weir, Philip J Guo, Robert C Miller, and Krzysztof Z Gajos. 2014b. Crowdsourcing step-by-step information extraction to enhance existing how-to videos. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 4017–4026.

15. Geza Kovacs. 2016. Effects of In-Video Quizzes on MOOC Lecture Viewing. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale (L@S '16)*. ACM, New York, NY, USA, 31–40. DOI: **http://dx.doi.org/10.1145/2876034.2876041**

16. So Young Kwon and Lauren Cifuentes. 2009. The comparative effect of individually-constructed vs. collaboratively-constructed computer-based concept maps. *Computers & Education* 52, 2 (2009), 365–375.

17. Justin Matejka, Tovi Grossman, and George Fitzmaurice. 2013. Swifter: Improved Online Video Scrubbing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1159–1168. DOI: **http://dx.doi.org/10.1145/2470654.2466149**

18. John R McClure, Brian Sonak, and Hoi K Suen. 1999. Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of research in science teaching* 36, 4 (1999), 475–492.

19. Toni-Jan Keith Palma Monserrat, Shengdong Zhao, Kevin McGee, and Anshul Vikram Pandey. 2013. NoteVideo: Facilitating Navigation of Blackboard-style Lecture Videos. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 1139–1148. DOI:**http://dx.doi.org/10.1145/2470654.2466147**

20. Joseph D Novak and Alberto J Cañas. 2008. The theory underlying concept maps and how to construct and use them. (2008).

21. Roni Reingold, Rikki Rimor, and Anat Kalay. 2008. Instructor's scaffolding in support of student's metacognition through a teacher education online course: a case study. *Journal of interactive online learning* 7, 2 (2008), 139–151.

22. Wolff-Michael Roth and Anita Roychoudhury. 1993. The concept map as a tool for the collaborative construction of knowledge: A microanalysis of high school physics students. *Journal of research in science teaching* 30, 5 (1993), 503–534.

23. M. Schwab, H. Strobelt, J. Tompkin, C. Fredericks, C. Huff, D. Higgins, A. Strezhnev, M. Komisarchik, G. King, and H. Pfister. 2017. booc.io: An Education System with Hierarchical Concept Maps. *IEEE Transactions on Visualization and Computer Graphics* PP, 99 (Jan 2017 2017), 1–1. DOI: **http://dx.doi.org/10.1109/TVCG.2016.2598518**

24. Ruey-Shiang Shaw. 2010. A study of learning performance of e-learning materials design with knowledge maps. *Computers & Education* 54, 1 (2010), 253–264.

25. Chao Boon Teo and Robert Kheng Leng Gay. 2006. A Knowledge-driven Model to Personalize e-Learning. *J. Educ. Resour. Comput.* 6, 1, Article 3 (March 2006). DOI:**http://dx.doi.org/10.1145/1217862.1217865**

26. Minhong Wang, Jun Peng, Bo Cheng, Hance Zhou, and Jie Liu. 2011. Knowledge visualization for self-regulated learning. *Educational Technology & Society* 14, 3 (2011), 28–42.

27. Sarah Weir, Juho Kim, Krzysztof Z Gajos, and Robert C Miller. 2015. Learnersourcing subgoal labels for how-to videos. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, 405–416.

28. Jacob Whitehill and Margo Seltzer. 2017. A Crowdsourcing Approach to Collecting Tutorial Videos–Toward Personalized Learning-at-Scale. In *Proceedings of the Fourth (2017) ACM Conference on Learning@ Scale*. ACM, 157–160.

29. Joseph Jay Williams, Juho Kim, Anna Rafferty, Samuel Maldonado, Krzysztof Z. Gajos, Walter S. Lasecki, and Neil Heffernan. 2016. AXIS: Generating Explanations at Scale with Learnersourcing and Machine Learning. In *Proceedings of the Third (2016) ACM Conference on Learning @ Scale (L@S '16)*. ACM, New York, NY, USA, 379–388. DOI: **http://dx.doi.org/10.1145/2876034.2876042**

30. Han Zhang. 2017. Smart Jump: Automated Navigation Suggestion for Videos in MOOCs. In *Proceedings of the 26th International Conference on World Wide Web Companion (WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 1183–1184. DOI: **http://dx.doi.org/10.1145/3041021.3055364**

31. Barry J Zimmerman and Dale H Schunk. 2001. *Self-regulated learning and academic achievement: Theoretical perspectives*. Routledge.

32. Krunoslav Zubrinic, Damir Kalpic, and Mario Milicevic. 2012. The automatic creation of concept maps from documents written using morphologically rich languages. *Expert systems with applications* 39, 16 (2012), 12709–12718.