

Computer-Assisted Learning Based on Universal Design, Multimodal Presentation and Textual Linkage

Leyla Zhuhadar¹  · Bryan Carson¹ ·
Jerry Daday¹ · Evelyn Thrasher¹ · Olfa Nasraoui²

Received: 11 October 2015 / Accepted: 8 March 2016 /
Published online: 31 March 2016
© Springer Science+Business Media New York 2016

Abstract Refining city services is gradually being placed in the hands of the citizens, or, as in the case of IBM's initiative, "let's build a planet of smarter cities" (<https://www-03.ibm.com/press/us/en/pressrelease/35573.wss>), at their fingertips. By reducing cost and gaining control in building smart transportation management systems, IBM provided a grant to the city of Chicago to reinvigorate Chicago's K-12 schools and city college. Similarly, Catherine Bracy (<https://www.codeforamerica.org/people/catherine-bracy/>) and her team at Code for America are using technology to "build governments for the people and by the people in the twenty-first century." It is evident that smart cities should accommodate every citizen, most especially those who may struggle with accessing information through conventional mediums. Specifically, smart cities must ensure that full participation is available to those citizens with communication and learning disabilities and new immigrant populations who experience difficulties understanding the language of their new home county. The biggest challenge for these

✉ Leyla Zhuhadar
leyla.zhuhadar@wku.edu

Bryan Carson
bryan.carson@wku.edu

Jerry Daday
jerry.daday@wku.edu

Evelyn Thrasher
evelyn.thrasher@wku.edu

Olfa Nasraoui
olfa.nasraoui@louisville.edu

¹ Department of Information Systems, Gordon Ford College of Business, Western Kentucky University, 1906 College Heights Blvd, Bowling Green, KY 42101, USA

² Department of Computer Engineering and Computer Science, University of Louisville, Louisville, KY 40292, USA

citizens is using the Internet for learning because of the need to read, view, or listen to content published online. Disabled students enrolled in science, technology, engineering, and mathematics (STEM) courses at colleges and universities are at a particular disadvantage since the use of empirical observation is critical in scientific learning and research. In this paper, we propose an intervention using a new software product and workflow for video captioning—a Universal Video Captioning platform (UVC). This platform provides a semi-automatic approach to synchronize the captioning into accessible STEM-related videos. It has the potential to transform learning and teaching for students with disabilities and those whose native language is not English by integrating synchronized captioned educational videos into undergraduate and graduate STEM disciplines. While accessing scientific content can pose unique challenges for disabled students and those who have just started learning a new language, the UVC platform has the potential to provide students with disabilities and recent immigrants the ability to pursue new and deeper learning opportunities.

Keywords Smart citizens · Smart cities · Universal design · Video captioning

Introduction

Today's smart cities use technology and systems to enhance interactivity by revamping infrastructure, management, communications, and transparency. The “Internet of Things” refers to connections between systems and appliances within our everyday environment, such as houses, vehicles, and office buildings. The Internet of Things can encompass the “1) Personal and Home; (2) Enterprise; (3) Utilities; and (4) Mobile” (Gubbi et al. 2013). Smart cities are one-step up from the Internet of Things, representing a “System of Systems” (Hernández-Muñoz et al. 2011). In effect, this is an aggregation of the Internet of Things, with systems talking to systems on a metropolitan level.

Smart systems and smart cities represent a host of possibilities, particularly for those with physical disabilities. They can also help to monitor people with a variety of medical conditions such as by monitoring pacemaker data of those with heart conditions (Hernández-Muñoz et al. 2011). Theoretically, a smart city would be able to help a deaf resident with real-time visual information, or a blind resident via Bluetooth audio. As Macagnano (2008) notes, “the disabled/elderly have only to synchronize onto a chosen menu (transport, shopping, health care, entertainment) and the surrounding environment becomes ‘really’ barrier-free, by offering visual/audio appropriate information. In this context, audible messages (Bluetooth, WI-FI and next generation wireless technologies) could also verbally describe to the user, in real time, what is happening (next bus stop name, name of shop owner, timetables, etc.) according to his/her 3D position in the city (Macagnano 2008).”

In working to improve outcomes for all in the smart city, we must consider the educational system. Smart schools will be both creators and consumers of data and information communication technologies. But, the pedagogical impact on students is foremost. Smartly interconnected systems and smartly interconnected cities will be for naught if no accommodations are made for educating students with communications or learning disabilities, or those whose native languages are not English. The K-12 and higher education systems are referred to collectively as P-20 education (Futrell 2014).

Science, technology, engineering, and math (STEM) students with disabilities are at a particular disadvantage since empirical observations are all vital to scientific research, discovery, and understanding. Captioning of audio and video lectures can help to alleviate these issues. The Universal Video Captioning platform (UVC) provides a semi-automatic approach to synchronize captioning into accessible STEM-related videos. It has the potential to transform learning and teaching for students with disabilities by integrating synchronized captioned educational videos into undergraduate and graduate STEM disciplines. It also takes full advantage of multiple learning styles, since research shows that students are more likely to learn the materials and to express satisfaction with courses that accommodate their learning styles. The Accessible Educational STEM Videos Project permits us to combine the advantages of captioned/transcribed multimedia with the recognized potency of audio and video to reinforce learning through the capability of re-listening or re-watching content. Multimedia objects with captioning may also assist students whose native language is not English, prior research describes this process in more detail (Zhuhadar 2015; Zhuhadar et al. 2007, 2009a, b, c, 2010, 2015; Zhuhadar and Nasraoui 2008a, b; Zhuhadar and Yang 2013). In 2009, 32.7 % of all US students in the STEM disciplines were foreign nationals (Wasem 2012). Use of subtitles is very common in foreign-language and English as a Second Language classes. Several authors (Cristea et al. 2006; Stewart et al. 2006; Stewart and Pertusa 2004) have noted that native English speakers studying Spanish who viewed films with the audio and subtitles in the same language learned more vocabulary, showed a gain in comprehension, and had higher levels of engagement and satisfaction with the material. Similarly, Hayati and Mohmedi (2011) reported a significant improvement among international students learning English, noting that multimodal delivery “[enhances] comprehension better than simply processing subtitles through silent reading” (p. 190). This line of research suggests that the use of captioned multimedia may also improve learning outcomes among STEM students whose native language is not English. The UVC’s significance will extend from the integration of multimodal delivery of audio and video in STEM courses to the transformation of curricular activities in P-20 education to accommodate students with disabilities.

Rational and Background

In recent years, P-20 education (pre-kindergarten through graduate education) has strived to serve students with disabilities more effectively. Information in Table 1 shows individuals with disabilities who are served by K-12 (kindergarten through secondary school) schools while information in Table 2 shows the relative distribution of students with disabilities in colleges and universities. In 2011–2012, K-12 schools enrolled 6,401,000 students with disabilities, while in 2013, higher education institutions enrolled 2,417,400 students who identified as having a disability.

Several statutes and standards require learning objects to be accessible to students with disabilities. These include the Individuals with Disabilities Education Act, 20 U.S.C. § 1400 et seq. (1990, amended 2004) the Rehabilitation Act (1973), the Americans with Disabilities Act (1990), and the World Wide Web Consortium’s Web Content Accessibility Guidelines (2008). Accessible learning, while often presented as an issue for distance education, is also a barrier for students with disabilities in face-to-face classes. In fact, distance education solutions hold some promise for helping

Table 1 Children 3 to 21 years old served under the Individuals with Disabilities Education Act, part B, by type of disability: school year 2011–2012

Type of disability	Number in thousands (percentage of total enrollment)
Total	6,401,000 (12.9 %)
Specific learning disabilities	2,303,000 (4.7 %)
Speech or language impairments	1,373,000 (2.8 %)
Intellectual disability	435,000 (.9 %)
Emotional disturbance	373,000 (.8 %)
Hearing impairments	78,000 (.2 %)
Orthopedic impairments	61,000 (.1 %)
Other health impairments ^a	743,000 (1.5 %)
Visual impairments	28,000 (.1 %)
Multiple disabilities	132,000 (.3 %)
Deaf-blindness	2000 (rounds to zero %)
Autism	455,000 (.9 %)
Traumatic brain injury	26,000 (.1 %)
Developmental delay	393,000 (.8 %)

^a Other health impairments include having limited strength, vitality, or alertness due to chronic or acute health problems such as a heart condition, tuberculosis, rheumatic fever, nephritis, asthma, sickle cell anemia, hemophilia, epilepsy, lead poisoning, leukemia, or diabetes

Source: U.S. Department of Education, National Center for Education Statistics. (2015), *Digest of Education Statistics 2013* (NCES 2015-011), chapter 2

students with disabilities enrolled in traditional face-to-face courses. Technologies used in distance learning courses can also assist non-native speakers struggling with the accessibility of information in coursework. To make online coursework accessible to students with disabilities, colleges and universities provide captioning services, which transcribe text from audio and video files to ensure that those with auditory disabilities access and learn content in courses. Moreover, figures, graphs, and other visual documents are tagged with explanations to ensure that those with auditory of visual impairments know the title and purpose of these images. The transcription of lectures for accessibility of disabled students in online courses could also benefit disabled students enrolled in face-to-face courses. However, transcription is a very labor-intensive and costly endeavor, and it poses an even greater challenge for disabled students participating in traditional face-to-face coursework, where new content is disseminated in lectures, labs, and discussion groups multiple times per week.

A review of the literature on learning shows that simultaneous multimedia experiences represent effective instructional design practices (González et al. 2015).

Metacognition is one important factor in the multimedia design puzzle. Cognitive learning theories emphasize the importance of context, allowing individuals the ability to store, process, and remember learned information in terms of pre-learning, learning, and performance (Tessmer et al. 1997, p. 90). Metacognition is the idea of an individual being cognizant of how he/she is learning. In simple terms, it is “thinking about thinking” and is a prerequisite for higher-order learning (Zohar 1999, p. 418). Given this knowledge, instructional designers and students need to take control of the

Table 2 Major field of study of undergraduates, by disability status: 2012

Major field of study	All undergraduates	Without disability	With disability
All fields (number)	21,823,100	19,405,700	2,417,400
All fields (%)	100	100	100
Non-S&E (%)	75.5	75.3	76.5
Business or management (%)	16.0	16.2	14.4
Education (%)	5.4	5.4	4.9
Health (%)	19.6	19.6	19.0
Humanities (%)	17.5	17.3	19.3
Other (%)	17.0	16.8	18.9
S&E (%)	24.6	24.7	23.3
Computer and information sciences (%)	4.3	4.3	4.8
Engineering (%)	5.0	5.1	4.3
Life sciences (%)	6.6	6.8	5.4
Mathematics (%)	0.5	0.5	0.3 ^a
Physical sciences (%)	0.9	1.0	0.9
Social and behavioral sciences (%)	7.2	7.1	7.7

Includes undergraduates with declared majors. For disability status, those who reported any type of disability related to blindness, deafness, severe vision or hearing impairment, substantial limitation of mobility, or any other physical, mental, or emotional condition that lasted 6 months or more were classified as with disability. Differences in the number of undergraduates in this table and tables 2-1–2-5 primarily reflect differences in the population of institutions surveyed. See U.S. Department of Education, National Center for Education Statistics, Estimating Undergraduate Enrollment in Postsecondary Education Using National Center for Education Statistics Data (NCES 2005-063) at <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2005063> for further explanation of these differences

S&E science and engineering

^a Interpret data with caution; the estimate is unstable because the standard error represents more than 30 % of the estimate

cognitive process, which requires an understanding of learning styles. These styles are an extension of personality and lead to very different ways of engaging with instructional material (Hauptman and Cohen 2011).

While assistive technology may help students with sensory or learning disabilities, the most important issue is proper pedagogical design (Moallem 2007–2008). As Thomas Klein (1995) pointed out in a much-cited article, “[C]omputers are not a panacea; they cannot rescue a school from weak teachers, a weak curriculum, or the absence of sufficient funding.” Universal design—making courses accessible for all students regardless of disabilities—goes beyond making courses usable and accessible. It involves “creating instructional goals, methods, materials, and assessments that work for everyone—not a single, one-size-fits-all solution but rather flexible approaches that can be customized and adjusted for individual needs” (CAST 2012).

Proper course design from the very beginning with attention to multimodal environments will help all students, not just those identified as having a disability (Moreno and Mayer 2007). As necessary accommodations vary by the type of disability, it can seem overwhelming to faculty members to design courses that meet these diversity of needs. For example, a blind student might be able to hear online videos but will miss

the visual details and a deaf student may be able to see the visual details but not listen to the professor's detailed explanation or elaboration of the visual artifact graphic. Generating simultaneous video captioning that describes the professor's explanation would be beneficial to this deaf student. However, a combined approach of describing the action and the dialogue will benefit all learners, especially those with visual or hearing impairments. Blind students can use a screen reader to listen to the captioning details, while deaf students will be able to read the transcription. Learning-disabled students will also benefit from the simultaneous multimodal presentation of information and materials helping to reinforce the material, concepts, theories, or research under investigation in the course.

Disabled students vary in their learning styles and in their type of required accommodation. Hearing-impaired students have a significant disadvantage in a traditional lecture classroom because of their inability to listen to formal lectures and explanations provided by the instructor. This is one of the primary motivating factors for online learning among students with hearing impairments as video captured lectures are transcribed for their use (Kisner 2015, p. 12). Likewise, multimedia captioning may also help visually impaired students allowing them to use screen reader technology.

The science, technology, engineering, and mathematics (STEM) disciplines are no exception in their need to serve students with disabilities (Jerome et al. 2008). The United States Federal Government Minority Science and Engineering Improvement Program. 20 U.S.C. § (1067) requires that universities make special efforts to recruit disabled students into these fields. STEM students with disabilities are especially impacted by the challenges they face, particularly given that color, sound, and observation are all important in science base courses and disciplines. For example, only one in four deaf students actually persist to graduation (Marschark et al. 2005). Technology may provide ways to overcome these challenges, particularly for students with sensory and learning difficulties using a diversity of multimodal and instructional design approaches. Given the importance of STEM disciplines to national competitiveness and future economic development, the US Government and private and non-profit entities have allocated resources specifically for the programs designed to promote the recruitment, education, and retention of students into STEM fields.

The emergence of distance education has resulted in innovative ways of teaching STEM programs using video capture, reusable multimedia, and interactive simulated labs. There are many advantages to adding captions, or transcriptions, to audio and video files. The capacity of multimedia to reinforce learning through re-listening or re-watching is powerful, and when captions are added to these multimedia files, the learning experience is more enhanced. These multimodal solutions help to make face-to-face, hybrid, and online classes and distance education available to all disabled students (Kidd 2009). Interactive learning that is "highly immersive" allows for the presentation of "three-dimensional objects" to teach relationships and higher-order thinking (p. 2109). This is especially useful in STEM disciplines where spatial intelligence and image processing are required (p. 2115). Students with cognitive-based impairments may also benefit from alternate delivery formats for the delivery of learning material (van Hoorebeek 2009, p. 232).

Use of subtitles is also very common in foreign-language and English as a Second Language (ESL) classes, suggesting that the use of captioned multimedia may also improve learning outcomes among STEM students whose native language is not English at English-speaking colleges and universities. Our global society requires smart cities to accommodate

those who come from other cultures and who speak different languages, including those who may know some English (or other languages at the center of instruction), but do not speak said language perfectly. The pre-kindergarten through secondary school -20 education system is no exception. In 2012, according to the National Center for Education Statistics, there were 4,397,318 English language learners in K-12 (kindergarten through the final year of high school, or secondary school) and 966,333 international students in American colleges and universities (National Center for Education Statistics 2015; U.S. Immigration and Customs Enforcement 2014, pp. 2.) The data further reveal that 344,299 (36 %) of international students are studying in the STEM fields (p. 23). Simultaneous multimedia learning may help all of these students.

There have been a number of studies of simultaneous multimedia use among language learners. Stewart and Pertusa (2004) noted that native English speakers studying Spanish who viewed films with the audio and subtitles in the same language learned more vocabulary, showed a gain in comprehension, and had higher levels of engagement and satisfaction with the material. Similarly, Hayati and Mohmedi (2011) reported a significant improvement among international students learning English, noting that multimodal delivery “[enhances] comprehension better than simply processing subtitle through silent reading” (p. 190).

In another student, Aldera and Mohsen (2013) randomly assigned 50 Arabic-speaking learners of English to three groups. One group learned with animation, captions, and keyword annotation (ACA). The second group used animation and captions (AC), while the control group received animation alone as the method of instruction (A). Participants were tested on vocabulary recognition, vocabulary production, listening comprehension, and listening recall. Tests were given immediately after the lessons, and again 4 weeks later. Interestingly, there were differences between the vocabulary and the listening comprehension/recall over time. Regarding vocabulary recognition and production, the ACA group had significantly higher scores than the AC group, whose scores in turn were significantly higher than the A group. However, the effect was reversed for listening comprehension and recall over time. On those tests, A had better scores than AC or ACA (p. 70). This research shows the importance of using multiple formats for the same material and of varying the approach throughout the course.

The importance of mindful (metacognitive) design cannot be overstated. With 12 % of the population identified as having disabilities (NSF 2013), and with international students making up over 4 % of the higher education community (Institute of International Education 2014), the need for well-designed learning objects will only increase. Instructional designers play an important role in ensuring that learning objects are accessible to all. Universal design and metacognitive practices hold the promise of increasing student learning for everyone. Students with sensory, communication, or learning disabilities and those whose native language is not English deserve the opportunity to learn at the same level as everyone else. Simultaneous captioning in multimedia provides the tools to make this happen.

Objectives and Novelty of the Universal Video Captioning Platform

The UVC platform will provide significant assistance to each type of disability listed in Table 2. Subsequently, alternative text such as captions or transcripts expands the

students' understanding of media. Since disabled students vary in their learning styles, some students may learn most effectively through oral means, while others may learn best through visual or textual means. Therefore, the significance of our approach is in providing a combined format where video is captioned and transcribed. This solution is significant because it holds the possibility of reaching students who have traditionally been at risk of not persisting. Research shows that the keys to serving disabled students are to make materials accessible in more than one medium and to ensure that “mental, physical, or other sensory disabilities” are considered when designing the learning experience (Richter and Paretti 2009; Seale 2013).

Why Universal Video Captioning Platform (Accessible STEM-Video)?

Automated closed captioning is considered a hard problem. Currently, there is no single product that provides a complete automated solution to this problem or that can accurately recognize a variety of voices instantly. Table 3 presents a comparison between most existing products and our proposed solution.

The infusion of video capturing with captions and/or transcripts will immerse students with disabilities in STEM programs through a multimodal presentation of materials so that they can interpret or reinterpret information in a more effective way. Students with disabilities will gain a better understanding of STEM materials through listening, watching, and/or reading, thereby broadening their understanding and deepening their skills. We anticipate that simultaneous multimodal presentations of materials will help disabled learners increase their grade point average and decrease their attrition. Our assumption is based on multiple studies that have found significant increases in grade point average and decreases in attrition among disabled learners who used assistive technology (Gutknecht 2015; Kuo and Kuo 2015; Owings et al. 2015; Raskind and Higgins 1999; Simoncelli 2010).

Proposed UVC Infrastructure Design

The emergence of the Web and new data-intensive computing platforms and technologies, such as WebDAV, Hadoop, MapReduce, Solr, and others, provide the best solutions to handle large amounts of data and the capability of searching for specific resources (Fig. 1). Therefore, we suggest a solution to this challenge where we can manage resources in a timely manner. The proposed infrastructure is a web-based platform that will be deployed on a server, allowing it to be accessed worldwide. It uses the latest WebDAV technology (Web-based Distributed Authoring and Versioning) to identify resources, users, content, etc. The front end will consist of the Solr Search Engine, Flex, and PHP, while the backend will be MySQL and Apache Tomcat servers.

Components of Platform Infrastructure

The platform infrastructure consists of an administrative management system, a faculty/staff user interface, a transcriber user interface, and a synchronized captioning applet.

Table 3 Comparison between most of the known products and UVC infrastructure

Software	Pros	Cons	Price	Can be used to add captioning to videos?
Dragon Naturally Speaking	<ul style="list-style-type: none"> •Speech recognition •Inexpensive 	<ul style="list-style-type: none"> •Recognizes only a single voice after a long period of training •Without training, the accuracy level is very low (around 20 %). •Does not recognize recorded audio and generally used for dictating speech 	\$99.99 per license	No: Can't be utilized on recorded audio or video
Camtasia Studio	<ul style="list-style-type: none"> •Inexpensive 	<ul style="list-style-type: none"> •Completely manual process •No voice recognition •User needs to add the script and time stamps for each word in the audio. •Time intensive •Desktop application 	\$299 per license	No: it is a desktop application. It Can't be applied on the server-side or be embedded as a Web-based service
3playmedia.com	<ul style="list-style-type: none"> •Web-based service •Full scripts and subtitle •Full synchronization process 	<ul style="list-style-type: none"> •Very expensive •Public universities cannot use their paid service 	\$3 per minute	No: Very expensive and in general, public universities cannot use their paid service
YouTube	<ul style="list-style-type: none"> •Free •Web-based service 	<ul style="list-style-type: none"> •Accuracy is very low ~20–30 % 	Free	No: Copyright issues
Open Source Speech Software from Carnegie Mellon University	<ul style="list-style-type: none"> •Free •Innovative solution •Research based desktop application •Open source •http://cmusphinx.sourceforge.net/ 	<ul style="list-style-type: none"> •Needs programming expertise •Desktop application •Does not provide a universal solution to all types of video formatting 	Free	No: Cannot be applied on the server-side but it could be embedded within our proposed platform
Universal Video Captioning Platform	<ul style="list-style-type: none"> •Free •Web-based service •Semi-automated process (i.e., the algorithm described in section c. will synchronize text with audio) •90–100 % synchronization accuracy 	<ul style="list-style-type: none"> •Not fully automated: however, an extension can fully automate the process by embedding the Open Source Speech Software from Carnegie Mellon University (http://cmusphinx.sourceforge.net/) 	Free	Yes

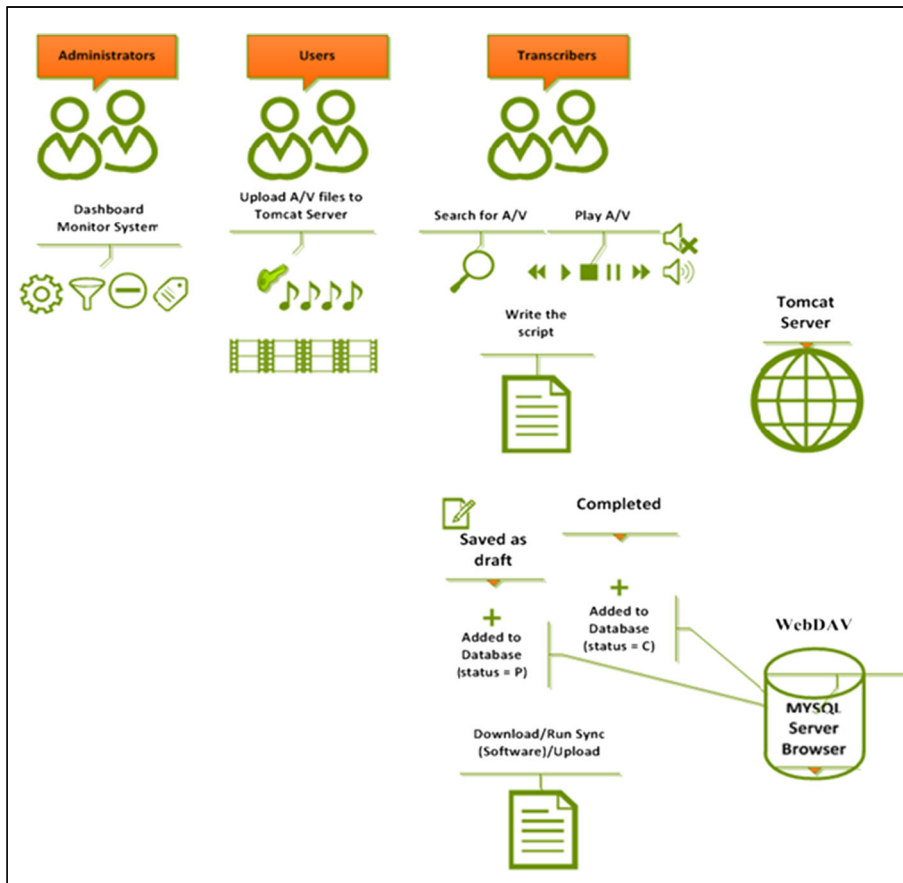


Fig. 1 UVC Infrastructure

The Admin Layer Interface

The purpose of the administration layer (Fig. 2) is to control user access, provide upload and download interfaces, maintain the database system, troubleshoot problems, and generate statistical reports. Finally, the administrative module provides a search engine mechanism to allow searching for a specific course, faculty member, video, caption, or script.

The User Layer Interface

The user layer interface module (Fig. 3) consists of the registration and login access. The uploading dashboard allows users to upload a video, an audio, or a reference to video or audio by using WebDAV technology. A universal identifier is associated with each resource. If the multimedia item is already located on the web, only a reference to that resource will be added to the database. The downloading dashboard lists the current status of each submitted request (download captioned video, download script, and status (“in process” or “completed”)).

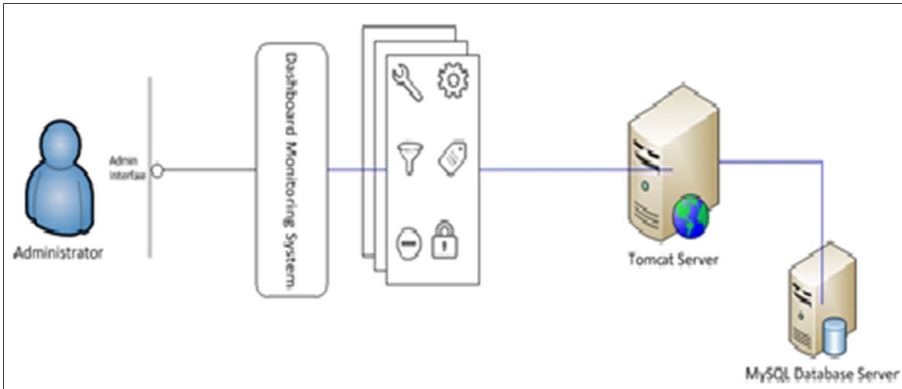


Fig. 2 Administration layer

The Transcriber Layer Interface

The transcriber layer (Fig. 4) provides registration and login access, as well as task management. Four dashboards are available for users. The transcribers will receive notices of requested audio or video files that need transcription. When a transcriber receives a notice, this task will be moved to his/her personal interface. The task will be removed from the queue, and the status of the resource will change to “in process.” A separate dashboard will list all of the queued resources on which the transcriber is working. Each transcriber will see a different (adaptive) interface listing the individual specific tasks. We are considering using open access to this platform and using crowdsourcing for transcription in the future. Once the project has been assigned, the transcriber will listen to the audio and transcribe on the platform in real-time. Scripts in progress can be saved as drafts. When the individual transcriber logs into the system again, all assigned and uncompleted work will show up. Once the work has been completed, the captioned video and script will be moved from the transcriber interface to the live user interface and the status of this specific task will change to “completed.” Instead of using a wave pedal as a hardware device, this functionality is provided through the use of arrow keys on the keyboard. An enhanced Web-based captioning

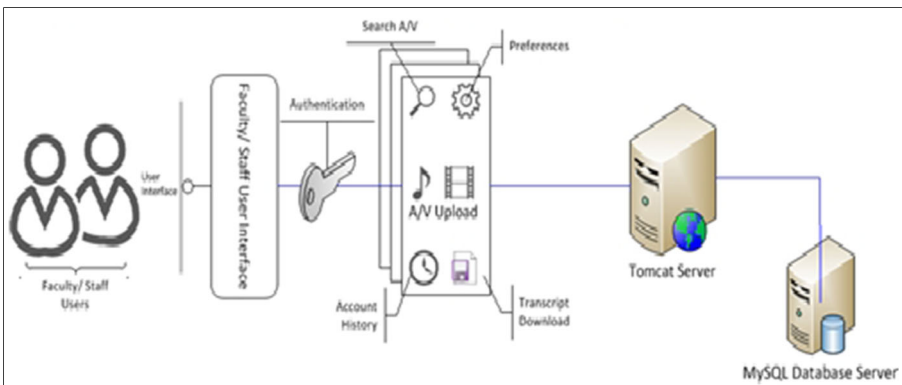


Fig. 3 User layer interface module

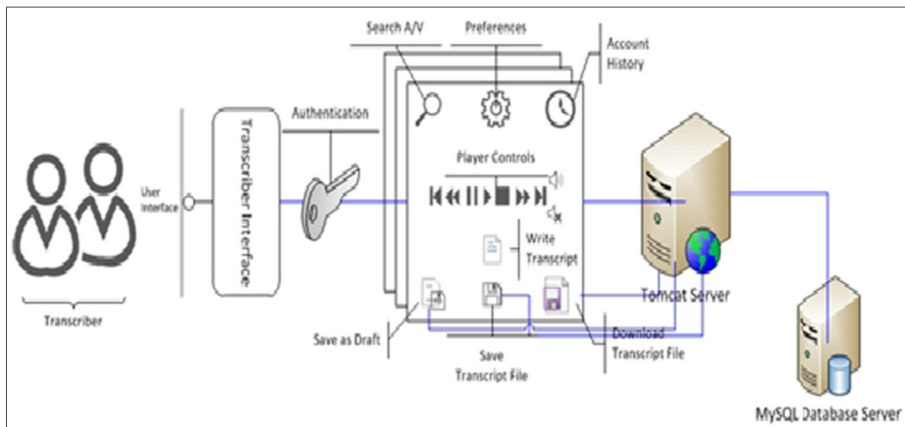


Fig. 4 Transcriber layer

Applet (Zhuhadar, L et al. 2015) allows transcribers to generate automatic synchronized captions with time stamps. Furthermore, transcripts with content description will be provided in plain text as additional content.

Synchronized Captioning Applet

As a starting point, we plan to develop an enhanced web-based version of our current software¹. We noticed that the software has three drawbacks: (1) it is not web-based and cannot be accessed in real-time over the web; (2) the accuracy of the synchronization between the audio and the script is affected negatively when there are long pauses; and (3) the application only works with specific video formats.

We therefore propose using a JavaScript Applet embedded inside the platform, which can be accessed easily by transcribers. To improve accuracy, we propose using a segmentation process to the audio file before we add the script. We will use short-time energy (Short_{time}En) to detect long pauses in the audio.

This measure can differentiate silence from speech and is calculated by the following equation [1]:

$$\text{Short}_{\text{time}}\text{En} = \sum_{-\infty}^{\infty} \left[x(m)w(n-m)^2 \right] = \sum_{m=n-N+1}^{m=n} x(m)^2 \quad (1)$$

Where $x(m)$ is audio signals, $w(m)$ is a rectangle window of length equal to N , and n is the index of the Short_{time}En. When Short_{time}En drops below a certain threshold, we consider this frame as a pause. After such a pause has been detected, we save its location. Therefore, there will be no captioning during long pauses. This process is fully automated since we are not going to make any physical changes to the original audio.

The speech that is converted to equivalent text in dashboard 2 (transcriber interface) will be selected, long pauses will be identified, text vocabularies will be divided over the length of the audio after extracting long pauses, and vocabularies will be embedded

¹ Temporarily omitted for reviewing purposes

in the correct time frame with no pauses. *As we know, the manual captioning is labor intensive. Therefore, this solution will have a significant impact on the system performance.*

Research Questions and Evaluation Methods

Our research questions involve the effectiveness of using captioning to improve outcomes, as follows:

1. *Does the integration of accessible STEM videos improve study skills for students with disabilities in STEM programs?*
2. *Does the integration of accessible STEM videos improve student grade point averages for students with disabilities in STEM programs?*
3. *Does the integration of accessible STEM videos increase student persistence in STEM programs among students with disabilities?*

Conclusion and Future Work

We presented a UVC infrastructure that provides a universal online model for video captioning to accommodate the unprecedented growth of video lectures. The UVC platform can serve as a repository for uploading videos and scripts. Captions can be automatically embedded inside educational videos using an enhanced and extended web-based version of our current software, to accommodate a variety of video formats. STEM teachers will be able to use the online UVC platform to upload audio and video and will receive a notification to download the original multimedia with synchronized captions and the transcript.

The proposed platform promises to (1) promote better opportunities for students with disabilities in STEM programs to understand their materials and (2) open new avenues for research on students' learning styles. The combination of online resources can foster our understanding of how students with disabilities benefit from the online learning and how they collaborate with their teachers.

After the implementation phase, we plan to measure the impact and effectiveness of adding captioning in educational videos on studying, grades, and persistence for individuals with disabilities compared to non-captioned videos (quasi-experiments). This will increase our understanding of students with disabilities in STEM programs and will shed light on their diverse talents and ways of learning. By transforming the online components of STEM education, we hope to help enhance career opportunities for students with disabilities.

References

- Aldera, A. S., & Mohsen, M. A. (2013). Annotations in captioned animation: effects on vocabulary learning and listening skills. *Computers & Education*, 68, 60–75. doi:10.1016/j.compedu.2013.04.018.
- Americans with disabilities act. (1990). 42 U.S.C. § 1201 et. seq

- CAST. (2012). *About UDL*. <http://cast.org/udl/index.html>.
- Cristea, A., Stewart, C., & Sirmakessis, S. (2006). *Adaptivity, personalization, and the semantic web*. New York: Paper presented at the APS '06: Proceedings of the joint international workshop on Adaptivity, personalization & the semantic web.
- Futrell, M. H. (2014). Transforming teacher education to reform America's P-20 education system. *Journal of Teacher Education*, 65, 271–283. <http://jte.sagepub.com/content/61/5/432>.
- González, M., Moreno, L., & Martínez, P. (2015). Approach design of an accessible media player. *Universal Access in the Information Society*, 14(1), 45–55.
- Gubbi, J. B., Rajkumar, M. S., & Palaniswami, M. (2013). Internet of Things (IoT): a vision, architectural elements, and future directions. *Future Generation Computer Systems*, 29(7), 1645–1660. doi:10.1016/j.future.2013.01.010.
- Gutknecht, E. A. (2015). *Exploring the effects of using the words their way program with teenage boys with autism*.
- Hauptman, H., & Cohen, A. (2011). The synergetic effect of learning styles on the interaction between virtual environments and the enhancement of spatial thinking. *Computers & Education*, 57(2011), 2106–2117. doi:10.1016/j.compedu.2011.05.008.
- Hayati, A., & Mohmedi, F. (2011). The effect of films with and without subtitles on listening comprehension of EFL learners. *British Journal of Educational Technology*, 42(1), 181–192.
- Hernández-Muñoz, J. M., et al. (2011). Smart cities at the forefront of the future Internet. In J. Domingue et al. (Eds.), *The future Internet: future internet assembly 2011: achievements and technological promises* [Lecture notes in computer science, v. 6656] (pp. 447–462). http://link.springer.com/chapter/10.1007/978-3-642-20898-0_32#.
- Individuals with disabilities education act. 20 U.S.C. § 1400 et seq. (1990, amended 2004)
- Institute of International Education. (2014). *Open doors*. <http://www.iie.org/Research-and-Publications/Open-Doors>.
- Jerome, M. K., Neuber, K., Stegall, B., Emenova, A., & Behrmann, M. (2008). Technology for integration of students with disabilities in higher education. *The engineering handbook of smart technology for aging, disability and independence*, 101
- Kidd, T. T. (2009). *Online education and adult learning: new frontiers for teaching practices*. IGI Global
- Kisner, J. L. (2015). *An examination of the experiences of deaf adult learners in online graduate academic programs*. Dissertations & Theses - Gradworks. <http://gradworks.umi.com/36/81/3681160.html>.
- Klein, T. (1995). Electronic revolution at the educational crossroads: foot-dragging on campus. *College Teaching*, 43(4), 151–155.
- Kuo, Y.-C., & Kuo, Y.-T. (2015). *Online learning and dropout prevention in K-12 education*. Paper presented at the Society for Information Technology & Teacher Education International Conference.
- Macagnano, E. V. (2008). *Intelligent urban environments: towards e-inclusion of the disabled and the aged in the design of a sustainable city of the future, a South African example*. Skiathos, Greece: 5th International Conference on Urban Regeneration and Sustainability: Sustainable City 2008. 24 - 26 September 2008. <http://hdl.handle.net/10204/3225>.
- Marschark, M., et al. (2005). Classroom interpreting and visual information processing in mainstream education for deaf students. *American Educational Research Journal*, 42(4), 727–761.
- Minority Science & Engineering Improvement Program. 20 U.S.C. § 1067(e)(1)(d) et seq
- Moallem, M. (2007-2008). Accommodating individual differences in the design of online learning environments: a comparative study. *Journal of Research on Technology in Education*, 40(2), 217–245.
- Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review*, 19(3), 309–326.
- National Center for Education Statistics, U.S. Department of Education. (2015). Table 204.20. Number and percentage of public school students participating in programs for English language learners, by state: selected years, 2002-03 through 2012-13. Local Education Agency Universe Survey, 2002-03 through 2012-13. https://nces.ed.gov/programs/digest/d14/tables/dt14_204.20.asp.
- National Science Foundation, National Center for Science and Engineering Statistics. (2013). *Women, minorities, and persons with disabilities in science and engineering: 2013*. [Special Report NSF 13-304]. <http://www.nsf.gov/statistics/wmpd/>.
- Owings, W. A., Kaplan, L. S., Khrabrova, I., & Chappell, S. (2015). Troops to teachers update changing, but still pleasing principals with high teaching quality. *NASSP Bulletin*, 99(1), 70–98.
- Raskind, M. H., & Higgins, E. L. (1999). Speaking to read: the effects of speech recognition technology on the reading and spelling performance of children with learning disabilities. *Annals of Dyslexia*, 49(1), 251–281.

- Rehabilitation Act of 1973, § 508, codified at 29 U.S.C. § 764(d) et seq
- Richter, D. M., & Paretti, M. C. (2009). Identifying barriers to and outcomes of interdisciplinarity in the engineering classroom. *European Journal of Engineering Education*, 34(1), 29–45.
- Seale, J. K. (2013). E-learning and disability in higher education: accessibility research and practice. 2nd edition. Abingdon UK: Routledge.
- Simoncelli, A. P. (2010). Designing online instruction for postsecondary students with learning disabilities. *Journal of Educational Multimedia and Hypermedia*, 19(2), 211–220.
- Stewart, M. A., & Pertusa, I. (2004). Gains to language learners from viewing target language closed-captioned films. *Foreign Language Annals*, 37(3), 438–442.
- Stewart, C., Cristea, A., Celik, I., & Ashman, H. (2006). *Interoperability between AEH user models*. New York, NY, USA: Paper presented at the APS '06: Proceedings of the joint international workshop on Adaptivity, personalization & the semantic web.
- Tessmer, M., & Richey, R. C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development*, 45(2), 85–115. <http://www.jstor.org/stable/30221388>.
- US Immigration and Customs Enforcement. (2014). *Student and exchange visitor information system: SVIS by the numbers, general summary quarterly review*. <http://www.ice.gov/doclib/sevis/pdf/by-the-numbers1.pdf>.
- van Hooerbeek, M. (2009). Analysing the parameters of the duty of reasonable adjustment to facilitate access to e-learning resources. *Multicultural Education & Technology Journal*, 3(3), 227–241. doi:10.1108/17504970910984899.
- Wasem, R. E. (2012). *Immigration of foreign nationals with science, technology, engineering, and mathematics (STEM) degrees*. Washington, DC: Congressional Research Service. <https://www.fas.org/sgp/crs/misc/R42530.pdf>.
- World Wide Web Consortium W3C. (2008). *Web content accessibility guidelines 2.0*. <http://www.w3.org/TR/WCAG20/>.
- Zhuhadar, L. (2015). A synergistic strategy for combining thesaurus-based and corpus-based approaches in building ontology for multilingual search engines. *Computers in Human Behavior*, 51, 1107–1115.
- Zhuhadar, L., Carson, B., Daday, J., & Nasraoui, O. (2015). *A Universal Design Infrastructure for Multimodal Presentation of Materials in STEM Programs: Universal Design*. Paper presented at the Proceedings of the 24th International Conference on World Wide Web, Florence, Italy.
- Zhuhadar, L., & Nasraoui, O. (2008a). Personalized Cluster-based Semantically Enriched Web Search for E-learning. Paper presented at the CIKM 08: ONISW The 2nd International workshop on Ontologies and Information Systems for the Semantic Web.
- Zhuhadar, L., & Nasraoui, O. (2008b). Semantic Information Retrieval for Personalized E-Learning. Tools with Artificial Intelligence, 2008. ICTAI '08. 20th IEEE International Conference on, 1, 364–368.
- Zhuhadar, L., Nasraoui, O., Wyatt, R. (2007). Knowledge Mining for Adaptive Multimedia Web-based Educational Platform. Technology Enhanced Learning: Best Practices. doi:10.4018/978-1-59904-600-6.ch010.
- Zhuhadar, L., Nasraoui, O., Wyatt, R. (2009a). Automated Discovery, Categorization and Retrieval of Personalized Semantically Enriched E-learning Resources. *International Conference on Semantic Computing*, 414–419.
- Zhuhadar, L., Nasraoui, O., Wyatt, R. (2009b). Dual Representation of the Semantic User Profile for Personalized Web Search in an Evolving Domain. Paper presented at the Proceedings of the AAAI 2009 Spring Symposium on Social Semantic Web, Where Web 2.0 meets Web 3.0.
- Zhuhadar, L., Nasraoui, O., Wyatt, R., Romero, E. (2009c). Multi-model ontology-based hybrid recommender system in e-learning domain. Paper presented at the Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology- Volume 03.
- Zhuhadar, L., Nasraoui, O., Wyatt, R., Yang, R. (2010). Visual knowledge representation of conceptual semantic networks. *Social Network Analysis and Mining*, 1–11.
- Zhuhadar, L., & Yang, R. (2013). The Impact of Social Multimedia Systems on Cyberlearners. *Computer in Human Behaviors*, 29(2), 378–385.
- Zohar, A. (1999). Teachers' metacognitive knowledge and instruction of higher order thinking. *Teaching and Teacher Education*, 15, 413–429. doi:10.1016/S0742-051X(98)00063-8.