Abstract

Syracuse University (SU) and Indiana University (IU), collaborating with the SU Bird Library, IU Wells Library and the Coulter Library of Onondaga Community College (OCC) in Syracuse, request $249,908 and will costshare $91,762 to support a National Digital Platform project. SU will serve as the lead applicant over the course of this three year project, July 1, 2018 through June 30, 2021.

Accessibility of digital collections has become a critical issue for academic libraries. Video captioning is a relied-upon method not only for providing essential accessibility for the deaf and hearing-impaired community—which is made up of 36 million Americans—but also for improving students' comprehension when consuming visual content. However, the cost of providing captions for a large volume of videos is prohibitive. As the National Association of the Deaf (NAD) has been taking legal action to require educational organizations to provide captioned online videos, many universities and community colleges are actively seeking cost-effective solutions to enable accessibility of video collections.

Researchers at SU's School of Information Studies and IU's School of Informatics and Computing, partnering with the SU Bird Library, the IU Wells Library, and the OCC Coulter Library, propose to 1) address the accessibility of video content by creating tools to provide a cost-efficient captioning service by applying Automatic Speech Recognition (ASR) and crowdsourcing closed captions; and 2) promote easy sharing, searching, and discovery of library resources by linking captioned videos with digital collections.

The team has extensive experience in relevant fields and propose a 2-phase plan. In phase one, tools as a platform-independent solution will be developed for creating machine generated captions both in real-time and offline. In phase two, accessibility metadata and user interaction designs will be created for captioned videos and integrated to link videos with other digital collections. The effectiveness of the tools will be evaluated in both phases, where usability study and performance tests will be conducted under different use case scenarios.

This project is a timely response to IMLS' National Leadership Grants for Libraries (NLG) *National Digital Platform* funding priority. By creating open-source tools, academic libraries nationwide will be able to help advance accessibility practices and improve the quality of service related to digital collections.

Narrative: Enabling Accessibility and Linking Digital Media Collections in Academic Libraries

Syracuse University (SU) and Indiana University (IU), collaborating with the SU Bird Library, IU Wells Library and the Coulter Library of Onondaga Community College (OCC) in Syracuse, request $249,908 to support the National Digital Grants for Libraries program through a project grant as part of the national digital platform category.

In this 3-year project, we propose to design, implement, and evaluate an open-source platform to enable accessibility and link digital media collections through partnership with academic libraries. Our proposed work will address the IMLS's National Leadership Grants for Libraries (NLG) *National Digital Platform* funding priority in that the deliverables can be used by academic libraries nationwide to help advance accessibility practices and to improve the quality of service related to digital collections.

Statement of National Need

Accessibility of digital collections has become a critical issue for academic libraries. As of 2012, it is estimated that 36 million Americans are hearing impaired (Blackwell, Lucas & Clarke 2014). The National Association of the Deaf (NAD) has been taking legal action to require educational organizations to provide captioned online videos to meet the needs of these individuals. In the meantime, there are approximately 1 million international students enrolled in undergraduate programs in the United States alone (Institute of International Education, 2015). Prior research has found positive effects for non-native English speakers when they used video captions as a tool for word, phrase, and context exposure (Chai & Erlam, 2008). However, even though there is a pressing need to provide accessible online videos for both hearing-impaired and non-native speaking users, a large proportion of videos are still not captioned. Both the SU and IU Libraries, for example, hold thousands of circulating visual media items in their digital collections along with streaming video content hosted on their university-owned sites. These media items and videos are not accessible to deaf and hearing-impaired users, because the majority of these video collections are not captioned.

One of the major reasons for the lack of captioned digital media and videos is the cost in time and money needed to create captions for these resources. According to research, it takes at least three to four times the length of a video for inexperienced people to create captions from scratch (Gaur, Lasecki, Metze & Bigham 2016), and experienced captionists are very costly. The rate can vary from $1/minute to $10/minute for professional closed captioning services. Machine-generated captions can reduce the labor cost (Munteanu, Baecker & Penn, 2008), however, captioning with Automatic Speech Recognition (ASR) technologies usually results in many errors (Hazen 2006). Another problem is that scaling up captioning services and costs are major obstacles for academic libraries in improving the accessibility to video collections. Captioning all video content being produced currently is an enormous task, even more so for the sizable number of legacy videos that are also in need of captions. Accessibility issues also occur when libraries do not have useful tools and/or services to offer to online course instructors, as a result, instructors may release their video lectures without captions.

Therefore, creating tools for enabling digital media collections' accessibility in academic libraries is crucial for realizing library services for all (Pionke, 2016), including people with disabilities and/or non-native English speakers who need caption assistance.

Additionally, without captions, searching and linking videos can rely only on metadata (Zeng & Qin 2015), e.g., video titles and brief descriptions. Providing captions for videos creates an opportunity for recommending and discovering relevant publications and other digital media collections with rich contextual information (Tuna, et al., 2017), e.g., full text transcripts together with video images. Thus, as

digital media repositories of all kinds increasingly become important venues for learning, research, and daily life information needs, accessibility enabling tools can play a key role not only for the deaf and hearing-impaired population, but also for linking digital media content to library resources and stimulate broader access to information resources.

In this 3-year project, we will collaborate with academic libraries, with guidance provided by an advisory board, to address the above need by designing and implementing tools as a cost-efficient captioning solution and to leverage captions for better linking to digital content in academic libraries. Presented below is our prior research relevant to the proposed work, our project design and evaluation plan as well as project team members and partners, who bring extensive expertise in the relevant fields and who are well prepared to conduct the proposed work.

Project Design and Evaluation

The overarching goal of this project is to enhance the accessibility of online videos and linkability of captioned videos with digital collections in academic libraries. According to the team's extensive experience with relevant prior work, we will achieve the goal through two phases. In Phase One, we will address the accessibility of video content by creating tools to provide a cost-efficient captioning service through the application of Automatic Speech Recognition (ASR) and by crowdsourcing closed captions. In Phase Two, we will promote easy sharing, searching, and discovery of library resources by linking captioned videos with digital collections. In both phases, we will evaluate the tools with different user scenarios at our collaborating academic libraries. A theoretical framework will be also used to guide our evaluation.

Prior Research on Video Accessibility

Given the high cost of captionists and the error-prone nature of machine-generated captions, researchers started taking a crowdsourcing approach to address the creation of captions (Kushalnagar, Lasecki & Bigham 2014). The idea is to have crowd workers edit machine-generated captions, which has been proven to be feasible. However, only using crowd workers who can hear well and who are skilled with English has likely left out potential contributors that are highly motivated. Prior approaches did not take into consideration that some crowd members who helped create captions for videos could be those who the captioning was intended to help most, e.g., people with hearing impairments and/or those who are learning the language (second-language speakers).

The PI's research team has developed a system, *BandCaption* (Huang et al., 2017), a web application that has already been deployed at SU. *BandCaption* structures the captioning tasks into a workflow that allows crowd workers with different language skills to help create captions at the micro-task level. More specifically, *BandCaption* starts with applying speech recognition technologies to create machine-generated captions first and then applies a Mark-Edit-Approve workflow to improve the quality of the captions. In brief, users with different needs can *mark* caption errors; users who have varying language skills can *edit* captions to correct the errors in a collaborative manner; and video owners can approve the caption edits.

We conducted a task-based study with 34 participants, including 4 hearing impaired students, 20 international students (with 10 students having low English proficiency and 10 with high English proficiency) and 10 native speakers to evaluate the performance of *BandCaption* (Huang et al., 2017). Our results showed that different user groups made complementary contributions based on their hearing or language strengths and constraints. For example, native speakers, in general, did best at marking all kinds of errors; hearing-impaired users were very sensitive to missing punctuation; and second-language

users picked up missing words and irrelevant words well, however they showed limitations in correcting the errors. In our study, native speakers could correct all the errors whereas second-language users failed to correct all errors because of complicated spelling or general unfamiliarity with spoken English.

Our study also showed that hearing-impaired users would send caption correction requests when they felt frustrated with caption errors (e.g., a long sentence without appropriate punctuation), and that displaying the number of correction requests could potentially encourage users to correct (more) caption errors as a result of altruism (Huang et al., 2017). International students were found to have difficulty in understanding fast speech and specific terms. According to our study with *BandCaption*, machine-generated captions can save about 94% of the typing effort compared to creating captions from scratch.

SU instructors and students now can access *BandCaption* through Blackboard, a widely used online content management system for teaching and learning. SU instructors and students use SU's Blackboard daily; this is where instructors and students can share video lectures and presentations, access other teaching materials, and submit homework assignments. Since it was released in Fall 2016, it has been used by 195 graduate students to watch video lectures. These included four courses, including Prof. Rachel Clarke' online classes on Cataloging of Information Resources and Information Resources: Organization and Access, Prof. Yang Wang's Data Visualization course, and Prof. Yun Huang's database management course.

However, *the current design of the BandCaption system has several limitations*. First, it can only work with YouTube videos, which are constrained by YouTube's ASR services and therefore only provides closed captions. Closed captions are a text version of the spoken part of a video. Second, the machine-generated captions need to be created before users watch the videos, and users cannot make real-time or offline requests for the captions. The system does not link videos with external online materials either.

Phase One – Accessibility

Building on the success of the PI's current research and understanding of the current system, the tasks during Phase One include devising a flexible and holistic solution to address the on-demand, real-time captioning needs for accessibility to digital video collections.

More specifically, the development work in the first phase will involve the following two major tasks. In *Task A1*, we will fill in web applications a void that captioning offline from a third-party cloud service is non-existent by creating a browser plugin and JavaScript libraries for machine-generated captions both in real-time and offline. We choose this approach because of its flexibility in working with any platform and their ability to display captions on top of the video without modifying the original video content. In *Task A2*, we will integrate crowdsourcing features of the existing *BandCaption* system with the browser plugin tool. These features for accessibility will be released as Application Programming Interfaces (APIs), such that they can be applied by other web applications.

Task A1: Creating a Browser Plugin and APIs for ASR-Generated Captions (in Real-time and Offline)

Prior research (Shiver & Wolfe, 2015) found that providing ASR-generated captions can improve deaf people's comprehension of video content, even though ASR-generated captions may contain errors. We have already identified a technically feasible and sound solution that can secure the delivery of machine-generated captions via a browser plugin.

One of the purposes of providing browser plugins is to serve real-time captions with videos. We will use a combination of open-source and third-party APIs on both the client and the server side. On the client side, the Web Audio API allows the audio stream to be captured from a video while the video is playing

(AnalyserNode, 2017). On the server side, the Google Speech API allows for a stream of audio to be continuously sent to the server and transcribed into text (Google Cloud Platform, 2017). By integrating both of these APIs, we will be able to perform real-time ASR and incorporate the results into our user interface. In addition to captioning in real-time based on streaming, our system could also extract audio content of the video in the background and transform it into captions by applying Google Synchronous Speech Recognition if the audio content is short or Google Asynchronous Speech Recognition if the audio is long (Google ASR, 2017). In general, the asynchronous service takes a long time to process the audio. This will ensure captions are generated regardless of whether the user watched the entire video or skipped parts, saving processing time in future requests.

The browser plugin will be released as an open-source project, and we will provide the APIs for other web developers to use the backend features for creating real-time and offline captions. As an illustration of how the APIs can be used, we will integrate the proposed APIs into the current *BandCaption* system and evaluate the system design in Phase Three.

Task A2: Integrating Crowdsourcing Features of BandCaption with the Browser Plugin Tool

Task A1 employs Automatic Speech Recognition (ASR) technology to bootstrap the system with machine-generated captions as a baseline. To this end, the baseline solution with ASR goes a long way towards addressing accessibility, however, corrections for ASR-generated captions may be requested for any particular video segment. Our approach allows for obtaining the initial captions and enhancing the quality of captions at a low cost. We will integrate the existing crowdsourcing components of *BandCaption* into the new browser plugin tool, such that the provided service can improve the accuracy of the machine-generated captions. For example, Figure 1 (a) shows the user interaction design for marking and editing captions of a particular video segment that is generated by ASR. One option to realize the crowdsourcing interaction could be to add a floating window with the BandCaption buttons and fields along the side of the current browser. Considering that users may display the video with full screen on, in *Task E1* (presented below), we will conduct user studies to examine the optimal user interface and interaction design of the tool.

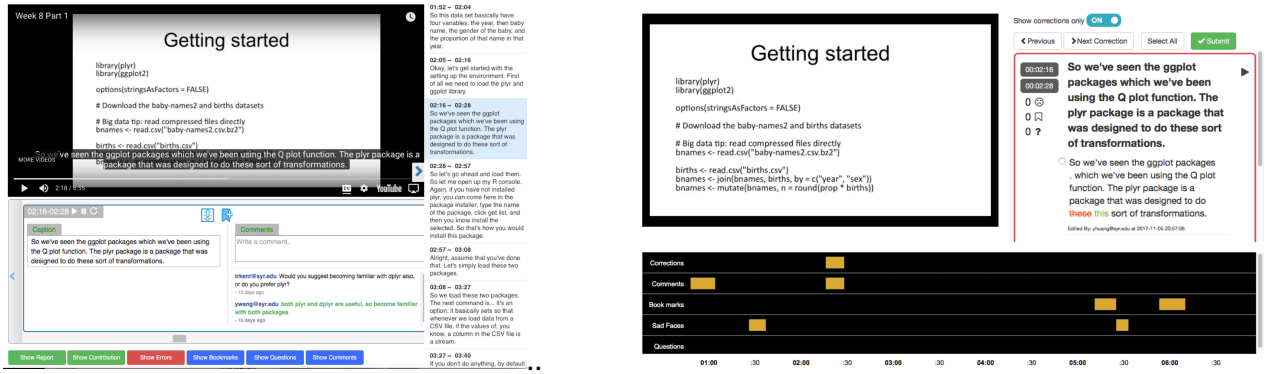

Figure 1. User interaction design: (a) video display panel where users can mark caption errors, edit captions, and navigate a video in a non-linear manner by clicking a caption in the transcript panel; (b) a user dashboard, where users can access bookmarks, comments, and other additional content that is linked to different segments of the video.

Phase Two - Linkability

Captions produced from Phase One offer a foundation for extracting descriptive and subject metadata and gathering referential metadata for richer content representation of videos. In Phase Two, the

captions will be archived to allow for offline access and the development of linking algorithms. More specifically, building on previous research in the structural, descriptive, and referential (SDR) metadata model and enhancing automatic metadata extraction with machine learning algorithms (Liu & Qin 2014), we will leverage video captions to develop metadata models for discovery of videos with accessibility and bidirectional links between captions and metadata for library resources. Linking between captions and wider library resources can be mutually beneficial: the use of captioned videos leads to more use of library collections while the accessibility metadata for videos incorporated into library search systems promotes the findability and usability of accessible videos. Below, we present details of the tasks to achieve the proposed goal.

Task L1: Understanding Prospective Users and Preparing Sources to Be Linked

In this task, we will conduct a survey and interviews with different stakeholders (faculty, individuals with/without hearing disabilities, and librarians) about linkability in order to gain insights into user needs when linking academic library resources with videos. Data collected from the survey and interviews will be analyzed to identify the characteristics of user needs. Based on the survey and interview results, we will also identify the preferred interaction with linked data under different use case scenarios, e.g., recommending library resources when online videos are being displayed or suggesting other caption-enabled videos while a user browses other types of digital collections.

Linking videos with captions to library collections or vice versa is essentially a matter of linking metadata between different systems. As library collections are increasingly becoming digital, whether they are created in house, acquired via purchase or subscription, or shared within a consortium, the most effective way to link captions with library collections would be through metadata APIs. Captions as a resource will be represented by metadata. The metadata representation will follow the principles of portability, extensibility, and linkability (Qin & Li 2013) to allow for automatic generation and interoperability with metadata in other systems. Such metadata representation will generate linked data sets for persons, organizations, events, time, geographical spaces, and relations among these entities. In particular, the metadata representation will incorporate accessibility metadata (Beyene, 2017) for resource discovery and use purposes by broader user populations. The accessibility metadata will inform information seekers whether the resource of interest (in this case, the videos) supports accessibility.

This task also involves understanding prospective users by identifying the sources that are potentially useful and linkable for this proposed project, which may be of three types: (a) independent digital libraries that archive digitized publications and provide metadata for discovery, (b) collections housed at academic libraries as represented by their catalogs, and (c) online videos that can be on the university websites or on the web outside of the university's. Because these types of sources may be managed under different business models and with different technical infrastructures, we will need to choose one or two examples from each type to test the tools we are developing.

Task L2: Developing Methods of Linking

We will first apply linked data technologies to link captions with library resources. This method involves several steps: (1) convert metadata records into a linked data format, the process of which may require indexing videos based on captions that are generated if metadata does not exist and creating metadata description triples in a linked data format; (2) establish graph representations for the linked metadata description (Liu et al., 2014); and (3) develop the linking mechanisms, e.g., text/graph mining and supervised topic modeling methods developed by the PIs (Jiang et al., 2015), for Linked Data to be connected with library search systems. Using Linked Data technologies will allow greater flexibility in linking between captions and library metadata, be it for a specific page, a paragraph in a publication or a

set of articles, books, or other types of resources. Eventually, various kinds of data will be integrated into a unique graph database (via Neo4j API) via data fusion methods (Liu, Jiang & Gao 2015; Jiang, Liu, Gao & Tang, 2016). Providing extracted metadata for captions in a Linked Data format enables the linked-out capability of captions to be extended to library search engines, and hence opens up a wide variety of channels from captions to open access repositories and library catalogs. Such linking mechanisms will be highly computational and can be automatically self-improved over time as more uses accumulate.

Another way of linking is to use Application Programming Interfaces (APIs) to link captions with library resources is an effective and straightforward method of linking. For example, our collaborator HTRC provides both bibliographic API and data API. The former allows one to access bibliographic and volume information (i.e., metadata), while the latter provides access to the actual data such as page images and OCR text in addition to associated metadata.

Evaluation

In both phases, we will conduct pilot studies to evaluate the effectiveness of the tools in turning audio into captions, the underlying system in supporting the crowdsourcing of caption corrections, and its linked data. As shown in Figure 2, we plan to study two scenarios: (a) users open a browser, possibly from a computer at the collaborating library where our browser plugin is installed, to watch an online video; and (b) users watch an online class video from the newly integrated *BandCaption* system with the proposed features. In either of the scenarios, both the accessibility and linkability features are integrated.

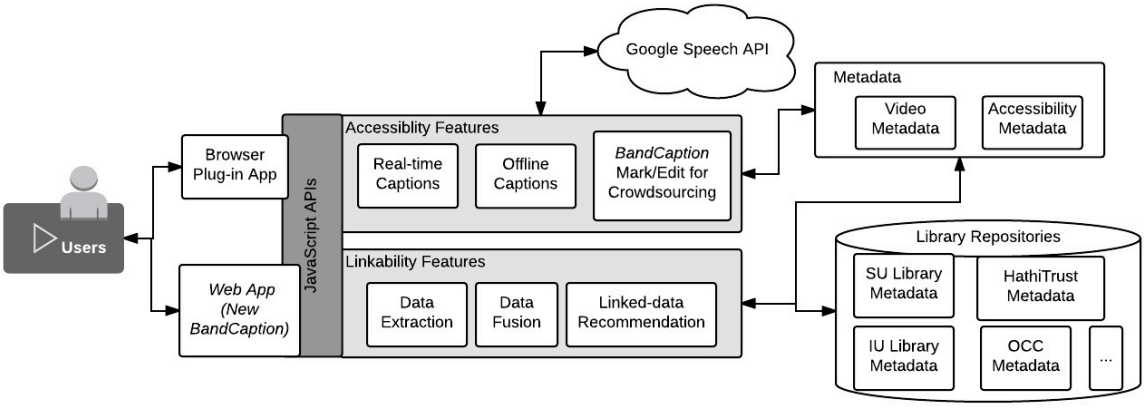

Figure 2. *Proposed System Model with Major Components*

The differences between the two scenarios may include, but are not limited to, the following: First, scenario (a) may be in a more informal learning environment, where users may randomly explore certain online content, rather than watching a class video in scenario (b). Second, within scenario (a) we may not be able to identify individual users since they do not need to log in using unique credentials, whereas users currently log in to BandCaption using their university account in scenario (b). Third, users in scenario (a) may have less incentive to help edit caption errors than those of scenario (b). Because students in scenario (b) have a stronger group identity (associated with the same course), this may trigger altruistic feelings and motivation to help their peers. These differences could impact the quality of the captions produced by *Task A2* and the execution of the linking algorithms by *Task L2.*

Task E1: Usability of Acquiring Captions and Motivation of Crowdsourcing Caption Edits for Accessibility

We will evaluate the usability of the interaction design of the proposed tools. Users should be able to send requests for captions and will be able to see captions generated by ASR. They should have the ability to edit the captions and see others' edits as well. Any actions taken (e.g., mark errors, edit captions) will be recorded and managed by our central server and become available to future viewers of the same video. We will conduct user studies to evaluate the system with hearing impaired users as well as international students whose native language is not English. We plan to use the Technology Acceptance Model (TAM) (Zimmerman et. al., 2011) to measure users' willingness to adopt the system powered by ASR and error corrections/edits enabled by crowdsourcing, thus uncovering the potential of the system in different circumstances. In both scenarios, we will explore more intrinsic motivating factors (e.g., curiosity, self-importance, etc.) and incentivizing mechanisms (gamification) to encourage users to correct caption errors (Kraut & Resnick, 2012) in different contexts.

Task E2: Evaluating and Optimizing Linkability

We will conduct task-based studies to evaluate the effectiveness of the model generated in *Task L2*, where digital library resources, e.g., books and papers, are recommended while users are watching videos. We will characterize users' needs through user feedback to linked data and by applying machine learning technologies (e.g., integrate various kinds of ranking functions via text/graph mining).

The team has extensive experience conducting this type of research. For instance, in Dr. Liu's prior work, a graduate-level information science course at Indiana University was used for running a preliminary experiment. Fifty-one students (masters and PhDs) voluntarily participated in this experiment (33 male and 18 female). They used a customized PDF reading system for 8 weeks, where relevant videos were recommended to them when they read papers using the system. The reading system was built upon the metadata of 41,378 publications (from ACM DL) and 786,555 videos (extracted from YouTube), and 9,263 keyword labeled topics were indexed in a heterogeneous graph for video recommendation. The preliminary evaluation showed that auto-recommended videos can provide important and useful information to help students better understand the content of the scientific publications, where 78.43% of students believe the recommended videos are very useful (Liu et al., 2015). Thus, linking between the captioned videos and digital collections can be optimized by characterizing users' personalized information needs based on their video-watching behavior. Without captions, only using video metadata, e.g., video text descriptions, could cause a bottleneck for the recommendation algorithm performance (Liu, 2013). Therefore, our proposed method of linking captioned videos with digital collections provides a key strategy to enhancing system performance.

We plan to conduct studies using the similar approaches to evaluate the performance of our linking solutions. In scenario (a), we will not have university student profiles for computational user profile generation, however, in scenario (b), we will be able to leverage student profile information. Methods such as sophisticated language model, graph-based random walk, plus supervised topic modeling proposed in the PIs' prior work (Jiang et al., 2015) will be employed for the optimization of the linking and recommendation algorithms. The indexed RDF and heterogeneous graph will be shared via an open-resource API.

Theoretical Framework for Evaluation

Since videos have become an increasingly popular medium for learning (Brandsteidl, et al. 2012; Defranceschi & Ronchetti 2011), we will also apply a theoretical framework—ICAP (interactive,

constructive, active, and passive) (Chi & Wylie, 2014) — to guide our evaluation. More specifically, this framework includes four progressive learning modes: passive, active, constructive and interactive. First, *Passive* is the lowest level of learning, where users receive the instructional information or materials without purposeful selection, discrimination, summarization, or extraction of information. For example, users only watching a video without taking any other action. Second, *Active* refers to the medium level of learning, where users have active interaction with learning materials with overt physical actions that can be observed apparently compared the cognitive activities in the mind. For example, users operate the video display by pausing, playing, fast-forwarding, and rewinding. Third, *Constructive* refers to learners' engagement with instructional materials accompanying externalized and self-generated outputs. For example, users add video captions. Fourth, *Interactive* is the highest level of learning in the ICAP framework, where students exchange instructional information or materials in the form of constructive learning and the process contains sufficient back-and-forth interactions. For example, students debate on the effectiveness of different links to video materials with their peers. We will collect and analyze system logs of user behaviors at different modes when conducting the proposed studies. Effective designs will lead to more user interactions at higher level of learning. For example, a design that engages users with more *Constructive* behaviors will be regarded more effective than a design that does not trigger any *Constructive* behaviors, if there is no difference on other behaviors.

Advisory Board Members

Partnering with academic libraries and experts in accessibility research is key for the proposed tasks to be successfully accomplished. The following scholars or practitioners from interdisciplinary backgrounds have agreed to serve on our advisory board and they provide complementary domain knowledge. We will consult with the advisory board members frequently to report our progress and seek feedback to review our results, as well as improve the tool design and efficiency of the solutions for accessibility and linkability.

- **Jeffrey Bigham,** Associate Professor in the Human-Computer Interaction Institute at Carnegie Mellon University. His work is at the intersection of human-computer interaction, human computation, and artificial intelligence, with a focus on developing innovative technology that serves people with disabilities in their everyday lives. Over the past few years, his group has been building and deploying crowd-powered technology to support blind people in their everyday lives.
- **Peter Blanck,** University Professor at Syracuse University and Chairman of the Burton Blatt Institute (BBI) at Syracuse University. The Burton Blatt Institute (BBI) reaches around the globe in its efforts to advance the civic, economic, and social participation of people with disabilities. Blanck has written over 200 articles and books about the American with Disabilities Act (ADA) and related laws. He has received millions of dollars in grants to study disability law and policy.
- **Robert MacDonald**, Associate Dean for Research and Technology Strategies of the Indiana University Libraries. Dr. MacDonald works to provide library information system services and discovery services to the entire IU system. In his role as Deputy Director of the Data to Insight Center, he works on new research related to large data analysis, storage and preservation through grant-funded and collaborative projects such as the HathiTrust Research Center. His research interests include technology management and integration of lean and agile frameworks, data preservation, learning eco-systems, and big data analytics.
- **Sharon M. Trerise,** Syracuse University IT Analyst – Accessibility, Information Technology Services, SU. Sharon's work focuses on promoting the implementation and use of accessible information and communication technologies (ICT). Her position involves developing a university-wide ICT Accessibility Policy, training staff and faculty in creating accessible course content, web pages and

documents, and evaluating University web content and new technology tools for accessibility.

Personnel and Partners

The team is well-qualified and prepared for the proposed activities. The project will be led by Dr. Yun Huang, PI, who has built and deployed several production systems and open software by collaborating with different stakeholders, including: Tiramisu (Huang et al. 2015; Tiramisu 2011; Tomasic et al. 2014; Zimmerman et al. 2011), Human Library (Huang, Debroski & Xia 2017; Jackson & Huang 2015; Web app 2014) and CAN (Composable Accessibility Infrastructure) (Huang & Dobreski 2015; Chrome 2014), etc. She has been awarded by NSF and Google as a single PI and DOE as a co-PI to conduct her systems research. Dr. Jian Qin, co-PI, will contribute her expertise in metadata and digital curation. Her work *Metadata 2nd ed*. (Zeng & Qin 2015) has just been named a 2017 CHOICE Outstanding Academic Title. Dr. Xiaozhong Liu, co-PI, will facilitate the design of knowledge-based video indexing, retrieval and recommendation (Liu et al., 2015; Jiang et al., 2016; Liu, 2013). The PI and co-PIs have been making successful collaborations on several research projects.

Dr. Xiaozhong Liu, from IU, has not received any IMLS award yet.

This project will enable a collaboration with the three libraries (from one public university, one private university, and one community college). It provides us a unique opportunity to study how size, private vs. public institution status, and policy factors inform and impact the design of the platform.

The project will be conducted in the PI's research lab (http://salt.ischool.syr.edu), where insights and feedback can be acquired from our lab members and advisory board members (who are prominent researchers) regarding a variety of topics, e.g., privacy, data security, and information policy, etc.

We will also partner with disability organizations that have a long history of research and community relationships with local disability communities (see the support letters from BBI and ARISE provided along with the proposal). These collaborations will provide an ideal platform to promote the adoption of the proposed technologies once it completes the prototyping phase and moves to the production phase.

Diversity Plan

Two of the team's PIs are female STEM faculty. The Ph.D. student who has already been involved in this project comes from an underrepresented group. He has extensive experience on web development and accessibility research. His work on developing and evaluating web and mobile systems for users with visual impairments was published at renowned conference venues such as UbiComp, W4A, ASSETS, and SOUPS. He has an award-winning record in international web design competitions.

We will continue encouraging women and minority students to participate in the proposed project activities. The SU iSchool has made great efforts to support minority groups, such as through BLISTS (Black and Latino Information Studies Support) and WIT (Woman in Technology). We will work with them in recruiting student researchers and ensuring equal participation in the project.

National Impact

This proposed project will provide a service platform that can be used by academic libraries nationwide in order to address the ever-expanding amount of visual content that needs to be made accessible in order to meet institutional and organizational policies and Americans with Disabilities Act (ADA) requirements. The infrastructure built from this project will also offer an opportunity to link videos with other digital collections, promoting access to library resources. This will benefit the library and scholarly community as a whole, but even more so, students, faculty and other users who are deaf or hard of hearing, as well as international students who are non-native speakers in English.

The three partnering libraries have different characteristics. SU Bird Library is operated in a mid-sized private university, IU Wells Library belongs to a large public university and the Coulter Library of Onondaga Community College (OCC) in Syracuse situates in a small-scaled community college. This is a great opportunity for us to understand different practices and create solutions that can be more adaptive and scalable. We also collaborate with disability organizations, i.e., ARISE and BBI, which have been serving the communities for years. They will help promote our work to better serve the disability communities.

The video accessibility problem is also prevalent outside of academia. Videos have become an increasingly popular medium for lifelong and informal learning. This has prompted the explosive growth of videos on social media and in online learning environments, e.g., Massive Open Online Courses (MOOCs), however, a majority of these videos are not captioned either. Commercial captioning services are expensive and often proprietary, requiring continued technical support with additional charges. Our tools can be used to promote accessibility of online videos in general.

Timeline

The schedule of completion presents our detailed 3-year action plan. Given the PI's prior research and system implementation experience, as well as the existing system and software management practice, the proposed work will be finished within the timeline. More specifically, our IRB applications for the proposed study have already been approved by SU and IU. Under the PIs' supervision, the graduate students will work together to implement the proposed tools and algorithms. Regarding the system implementation, the PI's team has developed different production systems, and released them on the market. The proposed action plan is a good estimation based on the past performance of the student team. We also have a well-established team programming environment in place, including version control software, a ticketing system, the protocol of server setup and a good practice of code review and release.

We will also provide comprehensive documents to support the continuing development of the system and its applications as an open sourced project on Github (https://github.com/). The SU iSchool has been providing server and storage resources to support the PI's research, which will ensure the continued availability of the tools once the project is ended.

Budget

The project budget will support graduate students' salary, PI and co-PI salary, dissemination travel, stipends for human subjects and advisory board members, a subcontract to Indiana University, and all associated fringe benefits and indirect costs. Additionally, Syracuse University will provide $90,206 in cost sharing to support the project. The budget justification document provides a detailed breakdown of costs. We also budgeted compensations for collaborators from OCC (the community college) and external advisory board members, who will provide feedback to work at major milestones and help recruit participants who will test tools.

Communications Plan

We will submit our work and findings to conferences such as ACRL and ASIS&T, as well as CSCW, CHI, and ASSETS where we believe our interdisciplinary research can draw a lot of attentions from researchers from other disciplines. A project website will be created to promote our project. Disseminating our findings is a critical step to share our experience and make a broader impact in the library and information science community as well as the STEM community, particularly in education.

Enabling Accessibility and Linking Digital Media Collections in Academic Libraries

*Yun Huang, *Jian Qin and Xiaozhong Liu
*Syracuse University and Indiana University

Schedule of Completion

Category	Task	Action Item	project year 1				project year 2				project year 3			
			Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
Phase One (Accessibility)	A1	Iterative design and implementation of a browser plugin for ASR-Generated captions (in real-time and offline)	X	X	X	X		X	X	X	X			
		Defining and releasing JavaScript APIs			X	X	X				X	X	X	
		Integrating the APIs in the existing BandCaption system, and revising interaction design	X	X	X	X	X					X	X	X
	A2	Integrating crowdsourcing features of the BandCaption in the browser plugin				X	X	X	X		X	X	X	X
		Releasing the APIs and writing reports/papers								X	X	X	X	X
Phase Two (Linkability)	L1	Developing interview and survey protocols	X				X	X			X	X		
		Running survey and interview studies		X	X				X	X		X	X	
		Collecting and analyzing user data			X	X	X			X	X		X	X
		Defining video and accessibility metadata				X	X	X	X	X	X			
		Extracting external resources' metadata								X	X	X	X	
	L2	Developing and optimizing linking algorithms		X	X	X	X	X	X	X	X	X		
		Linking data via data fusion methods			X	X	X	X	X	X	X	X	X	X
Evaluation	E1	Designing studies for scenario (a) and recruit participants	X	X	X	X			X	X	X			
		Conducting pilot studies and improving the interaction design				X	X				X	X		
		Analyzing the results and writing reports						X	X				X	X
	E2	Designing studies for scenario (b) and recruit participants		X	X	X				X	X			
		Running user studies and optimizing the linking algorithms				X	X				X	X	X	
		Analyzing the results and writing reports					X	X	X			X	X	X

DIGITAL PRODUCT FORM

Introduction
The Institute of Museum and Library Services (IMLS) is committed to expanding public access to federally funded digital products (i.e., digital content, resources, assets, software, and datasets). The products you create with IMLS funding require careful stewardship to protect and enhance their value, and they should be freely and readily available for use and re-use by libraries, archives, museums, and the public. However, applying these principles to the development and management of digital products can be challenging. Because technology is dynamic and because we do not want to inhibit innovation, we do not want to prescribe set standards and practices that could become quickly outdated. Instead, we ask that you answer questions that address specific aspects of creating and managing digital products. Like all components of your IMLS application, your answers will be used by IMLS staff and by expert peer reviewers to evaluate your application, and they will be important in determining whether your project will be funded.

Instructions
You must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

Instructions

☐ Please check here if you have reviewed Parts I, II, III, and IV below and you have determined that your proposal does NOT involve the creation of digital products (i.e., digital content, resources, assets, software, or datasets). You must still submit this Digital Product Form with your proposal even if you check this box, because this Digital Product Form is a Required Document.

If you ARE creating digital products, you must provide answers to the questions in Part I. In addition, you must also complete at least one of the subsequent sections. If you intend to create or collect digital content, resources, or assets, complete Part II. If you intend to develop software, complete Part III. If you intend to create a dataset, complete Part IV.

PART I: Intellectual Property Rights and Permissions

A.1 What will be the intellectual property status of the digital products (content, resources, assets, software, or datasets) you intend to create? Who will hold the copyright(s)? How will you explain property rights and permissions to potential users (for example, by assigning a non-restrictive license such as BSD, GNU, MIT, or Creative Commons to the product)? Explain and justify your licensing selections.

The computer source code will be open access for research purpose under the Creative Commons agreement. The computer source code will reside on a web server for free download.

A.2 What ownership rights will your organization assert over the new digital products and what conditions will you impose on access and use? Explain and justify any terms of access and conditions of use and detail how you will notify potential users about relevant terms or conditions.

The research team owns the copyright for the source code of the application.

The source code will be open access for research purpose by following the Creative Commons agreement.

The source code will reside on a web server for free download.

A.3 If you will create any products that may involve privacy concerns, require obtaining permissions or rights, or raise any cultural sensitivities, describe the issues and how you plan to address them.

We will notify the users that the system may collect user information for research purposes, and we will request users' consent before data collection.

Part II: Projects Creating or Collecting Digital Content, Resources, or Assets

A. Creating or Collecting New Digital Content, Resources, or Assets

A.1 Describe the digital content, resources, or assets you will create or collect, the quantities of each type, and format you will use.

Survey results will be text, and interview results can be both recorded in audios or noted in text.

System prototype design will be in figures.

Data sets resulted from assessment activities will be stored in SQL database and exported in CSV file format.

Reports and papers about the project are generated in Word, Latex and pdf format.

A.2 List the equipment, software, and supplies that you will use to create the content, resources, or assets, or the name of the service provider that will perform the work.

The project website and application servers will be hosted by the School of Information Studies at Syracuse University. Depending on the final decision choices, Drupal may be used for the creation and management of content hosted on the server.

A.3 List all the digital file formats (e.g., XML, TIFF, MPEG) you plan to use, along with the relevant information about the appropriate quality standards (e.g., resolution, sampling rate, or pixel dimensions).

Most of the design will be presented by JPEG figures, and we will try to provide high resolution.

B. Workflow and Asset Maintenance/Preservation

B.1 Describe your quality control plan (i.e., how you will monitor and evaluate your workflow and products).

Quality of interaction design will be controlled by walk through processes, where the researchers and designers will walk through the paper prototype of system design with different use case scenarios.

All content produced in this project will be documented with descriptive names, version numbers and dates.

Assessment data and analysis derivatives will be stored in a password protected server and accessible only by the PIs and IRB approved researchers.

B.2 Describe your plan for preserving and maintaining digital assets during and after the award period of performance. Your plan may address storage systems, shared repositories, technical documentation, migration planning, and commitment of organizational funding for these purposes. Please note: You may charge the federal award before closeout for the costs of publication or sharing of research results if the costs are not incurred during the period of performance of the federal award (see 2 C.F.R. § 200.461).

The source code and database will be stored on a permanent server maintained by the IT support department at the School of Information Studies. The maintenance cost of this server is covered by the School of Information Studies, Syracuse University.

C. Metadata

C.1 Describe how you will produce any and all technical, descriptive, administrative, or preservation metadata. Specify which standards you will use for the metadata structure (e.g., MARC, Dublin Core, Encoded Archival Description, PBCore, PREMIS) and metadata content (e.g., thesauri).

A self-developed data schema will be used to create database for hosting system and user data. It is possible to incorporate Dublin Core.

C.2 Explain your strategy for preserving and maintaining metadata created or collected during and after the award period of performance.

All metadata and data created during the project and after the grant period will be maintained on the servers hosted by the School of Information Studies at Syracuse University, and the maintenance will be operated by IT help support of iSchool at Syracuse University and the project PI.

C.3 Explain what metadata sharing and/or other strategies you will use to facilitate widespread discovery and use of the digital content, resources, or assets created during your project (e.g., an API [Application Programming Interface], contributions to a digital platform, or other ways you might enable batch queries and retrieval of metadata).

The project code and schema of the database will be available to anybody both on our project website and downloadable through GIT hub service.

D. Access and Use

D.1 Describe how you will make the digital content, resources, or assets available to the public. Include details such as the delivery strategy (e.g., openly available online, available to specified audiences) and underlying hardware/software platforms and infrastructure (e.g., specific digital repository software or leased services, accessibility via standard web browsers, requirements for special software tools in order to use the content).

The project code and schema of the database will be available to anybody both on our project website and downloadable through GIT hub service under the Creative Commons license Attribution-NonComerical-ShareAlike. We will also provide our contact emails on the sites, and anybody could send us an email to make their request.

D.2 Provide the name(s) and URL(s) (Uniform Resource Locator) for any examples of previous digital content, resources, or assets your organization has created.

http://salt.ischool.syr.edu

Part III. Projects Developing Software

A. General Information

A.1 Describe the software you intend to create, including a summary of the major functions it will perform and the intended primary audience(s) it will serve.

We will design, implement and evaluate an open-source platform to enable accessibility and link digital media collections through partnership with academic libraries and advisory board members. We will 1) address the accessibility of video content by creating tools that will provide a cost-efficient captioning service by applying Automatic Speech Recognition (ASR) and crowdsourcing

closed captions; and 2) promote easy sharing, searching, and discovery of library resources by linking captioned videos with digital collections. The deliverables can be used by academic libraries nationwide to help advance accessibility practices and to improve the quality of service related to digital collections. Patron users who have hearing impairments or who are non-native English speakers will also benefit from using the tools to access online videos by reading enclosed captions.

A.2 List other existing software that wholly or partially performs the same functions, and explain how the software you intend to create is different, and justify why those differences are significant and necessary.

This is an innovative and integrative tool for addressing accessibility of digital media collections. To realize part of the features, we will apply Google Speech APIs to create machine-generated captions. To the best of our knowledge, no existing software can wholly perform the same functions.

B. Technical Information

B.1 List the programming languages, platforms, software, or other applications you will use to create your software and explain why you chose them.

To implement the system, its web application and mobile applications, we plan to use the following: Java, SQL, Javascript, XML, JSON, and TOMCAT server, etc. They are widely applied programming languages and can be easily integrated with other open source tools.

B.2 Describe how the software you intend to create will extend or interoperate with relevant existing software.

The proposed system will use XML and JSON as input and output format, which facilitates convenient data management and transfer, enabling system interoperability.

B.3 Describe any underlying additional software or system dependencies necessary to run the software you intend to create.

Windows
Linux File systems

B.4 Describe the processes you will use for development, documentation, and for maintaining and updating documentation for users of the software.

In our previous projects, we have been using bitbucket, an integrated SCM & project management tool, to keep track of development progress, and tickets for handling different tasks. We will integrate Subversion, subversion control software to synchronize software development and system release. Project documents will be shared with and updated by team members on the bitbucket system.

B.5 Provide the name(s) and URL(s) for examples of any previous software your organization has created.

Bandcaption system (SU credential is needed to access the system) http://datascience.ischool.syr.edu
Developer Accessibility Tool https://chrome.google.com/webstore/detail/developer-accessibility-t/fcngonjbamdeknokaghjgdgkbgcdahdl?authuser=3
End User Tool - Enhance Web Page https://chrome.google.com/webstore/detail/enhance-web-page-

accessib/mceoeijcofnglhhnanfinbemamafooma?authuser=2

Human library http://humanlibrary.ischool.syr.edu

Mobile apps https://play.google.com/store/apps/details?id=edu.syr.ischool.humanlibrary

SU Indoor & Outdoor https://play.google.com/store/apps/details?id=edu.syr.ischool.orange.indoormap2&hl=en

The PI has also been working on the Tiramisu system http://www.tiramisutransit.com/ for almost 6 years, since she was a postdoc at Carnegie Mellon University. She is a co-owner of the company.

C. Access and Use

C.1 We expect applicants seeking federal funds for software to develop and release these products under open-source licenses to maximize access and promote reuse. What ownership rights will your organization assert over the software you intend to create, and what conditions will you impose on its access and use? Identify and explain the license under which you will release source code for the software you develop (e.g., BSD, GNU, or MIT software licenses). Explain and justify any prohibitive terms or conditions of use or access and detail how you will notify potential users about relevant terms and conditions.

The software code will be open-access for research purposes.

C.2 Describe how you will make the software and source code available to the public and/or its intended users.

We will create a project website to release the source code after project ends. We will also share the code on github.

C.3 Identify where you will deposit the source code for the software you intend to develop:

Name of publicly accessible source code repository: social computing systems lab site

URL: http://salt.ischool.syr.edu

Part IV: Projects Creating Datasets

A.1 Identify the type of data you plan to collect or generate, and the purpose or intended use to which you expect it to be put. Describe the method(s) you will use and the approximate dates or intervals at which you will collect or generate it.

The proposed system will design and implement tools to create machine-generated captions and to crowdsource enclosed captions for online videos of academic libraries. The system also links the videos with libraries' digital collections, such that users can access relevant digital collections more effectively. The data collection and generation can happen at a daily basis once the tool is released.

A.2 Does the proposed data collection or research activity require approval by any internal review panel or institutional review board (IRB)? If so, has the proposed research activity been approved? If not, what is your plan for securing approval?

Yes. We have sent our IRB applications, and they have been approved. We may need to modify the IRB applications when we identify it necessary. When the previous IRB applications are about to expire, we will renew the applications.

A.3 Will you collect any personally identifiable information (PII), confidential information (e.g., trade secrets), or proprietary information? If so, detail the specific steps you will take to protect such information while you prepare the data files for

public release (e.g., data anonymization, data suppression PII, or synthetic data).

Yes. It is possible the system collects community members' library IDs as PII. We will take the following steps to protect participants' identities during this research:
(1) All participant data will be stored based on numeric IDs generated for new users;
(2) the researchers will record any data collected during the research by the numeric IDs, not by their name or phone number; and
(3) all data will be stored on a server in a secure location that is accessible only by authorized researchers on this research team.

A.4 If you will collect additional documentation, such as consent agreements, along with the data, describe plans for preserving the documentation and ensuring that its relationship to the collected data is maintained.

We will collect consent forms at different tasks.
The first one is when we run interview studies. Consent forms are most likely to be printed out, and the signed copies will be collected and stored in PI's office.
The second one is when we run online survey or field trials. Consent forms will be displayed before participants start the studies. These consent forms will be stored on the study servers, and kept separately from the user data.

A.5 What methods will you use to collect or generate the data? Provide details about any technical requirements or dependencies that would be necessary for understanding, retrieving, displaying, or processing the dataset(s).

All user generated data will be collected via app interfaces and transferred to our server using https protocol. There will be different levels of access controls defined for each types of users, such that data will be retrieved, displayed and processed in a secure and privacy-preserving manner.

A.6 What documentation (e.g., data documentation, codebooks) will you capture or create along with the dataset(s)? Where will the documentation be stored and in what format(s)? How will you permanently associate and manage the documentation with the dataset(s) it describes?

We will create our own data model. During the project, user data and software code will be stored on the iSchool server dedicated for research. They are stored in scripts, txt, CSV, audio or video formats. The iSchool ITS conducts regular backup for the entire server. All data is stored on disk, all disks are backed up nightly.

A.7 What is your plan for archiving, managing, and disseminating data after the completion of the award-funded project?

After the completion of the research activity, user data and software code will be kept on the iSchool server dedicated for research. The SU ITS conducts regular backup for the entire server. All data is stored on disk, all disks are backed up nightly. The maintenance cost of this server is covered by the School of Information Studies at Syracuse University.

A.8 Identify where you will deposit the dataset(s):

Name of repository: We will create a website for this project, and the link will be shown on our center's website.

URL: http://salt.ischool.syr.edu

A.9 When and how frequently will you review this data management plan? How will the implementation be monitored?

The data management plan will be reviewed upon completion of each task of the project.

We will release data that can be publicly available. We will make sure we keep our partner libraries and advisory board members updated with our data management plan.