**Type of Contribution: PAPER**

**Digital Social Services: From Data Aggregation to Culturally Competent Content**

**Dan Albertson,** University at Buffalo, The State University of New York, United States, dalbert@buffalo.edu

**Keywords: Digital social services, cultural competence, data, online content**

**Introduction**

Digital social services utilize computing technologies to provide public outreach and/or support. These projects are many times funded by local, state, and/or federal government agencies in order to help address a particular problem, discrepancy, or need within the community. In order to serve communities, digital social services must be able to reach, communicate, and be used effectively by diverse populations.

The need and requirements for cultural competence (in addition to accessibility) are becoming more and more recognized (Ilomäki et al., 2016). Culturally competent information can be interpreted effectively and avoid misuse or misinterpretation, including across different cultures, languages, and/or backgrounds of individuals. Cultural competency is imperative across a variety of information services (e.g. governmental, health care, education) and modes of communication, increasingly in digital contexts where considerations for data processing and management are needed (Chang et al., 2004).

Cultural competency within digital social services is the topic presented; a particular project is used as a case to describe different types of data tasks, content needs, and roles of information professionals. While the present paper reflects on more practical aspects of data work, future data-driven research streams will also emerge from the project, which are also described here.

Data Tasks and Content Needs

Efforts of an ongoing project have resulted in specific types of data tasks and requirements for culturally competent content, a necessity of an active digital social services project.

*Project Overview*

The project *Multimedia Peer-to-Peer Abuse Prevention* has been funded by a state agency (within the US) to develop a digital library and online peer-to-peer support forum that will provide educational information, resources, and other digital tools that support individuals with intellectual and developmental disabilities who are experiencing various forms of exploitation and abuse. High quality information resources are being identified that will help individuals build personal capacity. The resources and tools developed through this project will be accessible on both standard computers and mobile devices, and will be fully accessible to individuals with different abilities and from diverse cultural and linguistic backgrounds. Digital social services with online support mechanisms can provide a number of benefits to users including a sense of empowerment, reduced loneliness, improved knowledge, decision-making skills, social well-being, and optimism (Barak et al., 2008; van Uden-Kraan et al., 2009). However, a number of tasks and considerations are necessary for ultimately achieving culturally competency for digital content.

*Content Aggregation and Data Processing*

Many times, the emphasis of digital social services is on the aggregation of quality content from different sources (as opposed to content creation). The primary reason for aggregating content is that it is difficult for people who need assistance from social services to find the information they need, as resources are often siloed, the authority and significance of the information is unclear, the information is not objective, resources may be out of date and/or incomplete, or users may not feel confident in their abilities to search and find information online (Albertson and Ju, 2016; Amsbary and Powell, 2003; Flanigan and Metzger, 2007; Metzger, 2007).

One primary objective for content aggregation is to collect diverse types or formats of information from high quality sources, with particular emphasis on multimedia so that users with different abilities and literacies can use and understand the content. A vetting system needs to be in place when aggregating content, particularly considering its use by potentially vulnerable audiences. Ultimately, for digital social services, all resources and content to be provided – regardless of information type – should adhere to standards of cultural competency, which will require a number of different processes.

First, data from aggregated content must be extracted and processed. Textual data can be extracted from source files of textual documents and from time-based media, i.e. videos, using closed captioning or automatic speech recognition (ASR) output. Once full text has been extracted, additional layers of processing can be applied to initiate progress toward cultural competency, such as parsing textual documents and formatting transcripts in structures compatible with streaming video and players.

Next, translation of all data – regardless of format – will be a major step in data processing. To prioritize which languages, many times those most commonly spoken in particular areas have been identified through state or local surveys and published in publically accessible data. Such public data can serve as indicators for which languages to prioritize for the content of digital social services.

Most importantly, decisions will have to be made on the translation approach based on the number of languages and the volume of content. Further, there is an obvious tension between automated translation approaches versus manual, particularly when a sizable collection is involved. Automated translation approaches are available through use of programming APIs, which offer speed and efficiency for processing large volumes of data into many different languages, with modest accuracy in syntax. On the other hand, manual translation (particularly that which is “certified”) offers higher levels of semantic accuracy. To summarize, automated translation runs a greater risk of miscommunicating or misunderstanding the semantics of messages, while manual or certified translation of even a modest dataset into several languages will result in a substantial cost. Therefore, while ideally all translated content provided by digital social services would be certified, a balance between what can be systematically translated versus that which cannot needs to be found.

There are other factors which are significant for data aggregation and processing. These include ensuring data extraction and parsing produce low-level content, or that with basic language structures, information that is concrete or factual, with limited abstraction and/or required interpretation on behalf of the user, preservable document structures which are compatible with different writing systems, and other important aspects of digital literacy. Effort given during the vetting process (when aggregating resources) will in turn aid in these qualities.

To summarize, the primary steps to assist with cultural competency in digital social services include:

* Aggregating multimedia (as a primary resource) and content vetting
* Data extraction directly from source files and content
* Translation and other data processing
* Testing content in user interfaces and structured documents

Once these steps of data aggregation and processing have been conducted, whether for textual or multimedia, appropriate techniques for organizing, indexing, or retrieving information can be tested and applied to enable access and classification of resources; these aspects of the project twill be reported on at a later date.

Opportunities **for Information Professionals and LIS Education**

With the constant move to online, information professionals are prepared to contribute to and work for such public and social services in a number of ways. In addition to data and user-centered design skills, information professionals are many times provided with experience in interfacing with the public as part of their education and/or on the job training. Thus, information professionals directly experience the need and importance of cultural competency on a regular basis. Data and public service skills coalesce nicely into an applicable skill set for digital social services. As a result, there are opportunities to include courses within degree programs that examine both digital data (processing and management) skills and cultural competency – together – which would be a unique contribution of an information studies degree (Ilomäki et al., 2016; Chang et al., 2004).

**Future Work and Research Streams**

*Evaluation*

Future activities on the project will include evaluations to measure the effectiveness of the digital tools and resources, particularly as they comply or adhere to cultural competency. The planned evaluations will be useful for future digital library projects, which traditionally emphasize usability and other quantitative assessments. While usability is critical (and will be part of the ongoing evaluation measures of the current project) the work as presented here demonstrates the need to also measure other issues and requirements of digital social services as they relate to cultural competency (e.g trustworthiness, confidence, confidentiality, and sensitivity) from a qualitative approach. This work provides a new perspective of digital library evaluation and assessment in showing how digital projects can benefit from qualitative evaluations of use, which has been mostly assessed through quantitative measures up until now.

*Research*

Different data-focused research streams will emerge from the project. Curation and culturally competent content, as examined together, can comprise a notable research area. Here, researchers can examine how informational lessons or topics can be organized and guided to ultimately support and improve learning across a wide range of users from different culturally backgrounds and abilities. Additionally, data ethics can be examined including topics of confidentiality, user safety, content moderation of user contributed information on sensitive topics, and on encouraging user buy-in for sharing personal details of lives.

REFERENCES**:**

Albertson, Dan, and Ju, Boryung. 2016. "Perceived self-efficacy and interactive video retrieval." *Journal Of Documentation* 72, no. 5: 832-857.

Barak, Azy, Boniel-Nissim, Meyran, and Suler, John. 2008. "Fostering empowerment in online support groups." *Computers In Human Behavior* 24, no. 5: 1867-1883.

Chang, Betty L., Bakken, Suzanne, Brown, S. Scott, Houston, Thomas K., Kreps, Gary L., Kukafka, Rita, Safran, Charles, and Stavri, P. Zoe. 2004. "Bridging the digital divide: Reaching vulnerable populations." *Journal of the American Medical Informatics Association* 11, no. 6: 448-457.

Flanagin, Andrew J., and Metzger, Miriam J. 2007. "The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information." *New Media & Society* 9, no. 2: 319-342.

Ilomäki, Liisa, Paavola, Sami, Lakkala, Minna, and Kantosalo, Anna. 2014. "Digital competence – An emergent boundary concept for policy and educational research." *Education and Information Technologies* 21, no. 3: 655-679.

Metzger, Miriam J. 2007. "Making sense of credibility on the Web: Models for evaluating online information and recommendations for future research.” *Journal of the American Society for Information Science and Technology* 58, no. 13: 2078-2091.

van Uden-Kraan, C.F., Drossaert, C.H.C., Taal, E., Seydel, E.R., and van de Laar, M.A.F.J. 2009. "Participation in online patient support groups endorses patients’ empowerment." *Patient Education and Counseling* 74, no. 1: 61-69.

ACKNOWLEDGEMENTS

The authors of the paper wish to thank the New York State Developmental Disabilities Planning Council for financial support.