Preserving Virtual Worlds Educational Events using Social Media Networks and Cloud Storage Services

Patricia C. Franks & A. Marie Vans; San José State University; San José, CA, U.S.A.

Abstract

Documenting and preserving educational experiences in virtual worlds can be a challenge, especially for semester-length courses of up to 16 weeks. Vast amounts of data are generated such as objects created in-world, project data contained on social media networks, and real-time educational events. While it is probably not feasible to preserve all the related data for reuse in later courses, it is possible to document essential data for later development of similar courses. Artifacts connect past educational experience with future design of educational material for distance education in Virtual Worlds. These courses were taught on the Second Life School of Information Island and sponsored by the School of Information at San José State University. We describe our approach to archiving this data, which is spread across the Internet using cloud storage services.

Introduction

"Virtual Worlds are History," as Henry Lowood has recently pointed out [1]. This rings true for instructors and developers involved in creating educational experiences in Virtual Worlds. At the San José State University iSchool, several courses have been run entirely in virtual worlds created to reflect a specific historical time period. As these historical time periods change with the start of a new course, the old worlds are taken down to make room for new worlds. In effect, these older worlds are history in the first sense of Lowood's phrase, in that they disappear once a new world is created in the same space. However, while the worlds are manifest and used as classrooms, these worlds are history in the second sense, that they have a history of their own in terms of events and activities that occur within them, and some record of these should be preserved to help creators and instructors learn to develop their own courses from prior experiences and successes.

Virtual worlds are frequently compared to games and many have used preservation practices when attempting to define the appropriate methods for preserving virtual worlds [2, 3, 4]. Some researchers have created archival information packages (AIP) consistent with the Open Archival Information System (OAIS) for preserving game objects with success for well-defined games with expected interaction between players [5]. While some of these approaches make sense for preserving educational virtual worlds, the investment in time and cost is prohibitive for our purposes. Another approach includes ideas from digital curation of 3D CAD models [6]. The challenges for preserving 3D objects are similar to the challenges for virtual world objects; however, we are interested in documenting experiences as well as the in-world objects.

In fact, because preservation of virtual worlds is so difficult, our goals are geared more toward documentation rather than preservation.

The "Living in Tudor Times" section of Virtual Worlds: Traveling through Time and Space was taught in-world in Second Life (secondlife.com) in fall 2013, and this was the first time such a course was offered at San José State University (Figure 1). The goal of this 16-week course was to teach Principals of Instructional Design using Discovery Learning [7], Constructivism [8], Social Development Theory [9], and Communities of Practice [10] utilizing a virtual world distance-learning platform. These goals were realized through a series of in-world projects, events, presentations, and the use of social networking sites for student assignments. The course has been offered since fall 2013, based on different themes that required distinct builds for each course, including Renaissance Italy (Figure 2) and Pre-Revolutionary France (Figure 3). Construction of the U.S. Revolutionary War site is underway for the summer 2015 term. Students are required to assume a character from the particular time frame represented during the semester. They create their living space in accordance with their societal station and then provide tours of their homes. They attend in-world lectures and present on various aspects of life during that time. Several events, such as a Tudor Reunion (Figure 1) in which students were required to appear "in character" and field trips to similar sites within Second Life occurred during the semester. A major assignment required the students to create a machinima documenting information about their character, providing a tour of their home and a discussion of a few important aspects of culture. The final project required a presentation and creation of an exhibit in-world, based on in-depth research into some aspect of life during the time period represented (Figure 2). During the semester, students utilized the social-networking platform, Spruz (spruz.com) to submit papers and discussion assignments. A vast amount of data, both textual and visual was being generated and it was spread across the Internet. What is needed is an implementation plan for archiving this data for future use.

Figure 1. Tudor Times (Fall, 2013) Top left: Castle. Top right: part of the village. Bottom left: Dance event. Bottom right: Meeting with Henry VIII.

Figure 2. Renaissance, Italy (Spring, 2014) Top left: Class in front of the Palazzo Farnese. Top right: Student house in village. Bottom: Exhibit in Renaissance Fair of the building of the dome in the Florence Cathedral.

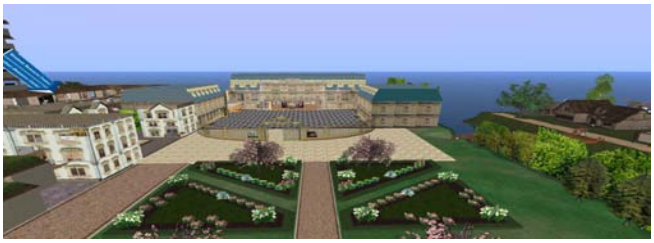

Figure 3. Prerevolutionary, France (Fall, 2014) Top: Build of the Palace of Versailles and village. Bottom right: Event dance in the Hall of Mirrors. Bottom left: Exhibit in Fair.

Utilizing Social Networks for Archiving Data

There has been interest in the use of social networks for preserving and archival purposes [11], with some studies focusing on a single platform, such as Pinterest [12], Instagram [13] and Flickr [14]. The question we are interested in, however, is how to manage several different social networks as a cohesive package, as we utilize several different sites during a single semester.

Each semester, a new site is created in Spruz and during the course of the semester, many of the educational experiences are recorded by students reflecting on their experiences, submitting assignments, and posting their own images. Machinima (video recording of in-world events and activities) and images are good ways to document events in virtual worlds [15], albeit from a single perspective--that of the person behind the video recording or taking photographs. Several videos have been created and are currently stored on a YouTube account (Figure 4) created specifically for documenting important events during the semester including student projects and presentations as well as invited speakers. This channel is currently located at:

www.youtube.com/channel/UCe3kkzyWCmzLGBXulzn3mzg

Other social networking accounts opened include a Pinterest account (pinterest.com/sjsuischool/virtual-worlds/)(Figure 5) to hold images and a WordPress blog (Figure 6) (ischoolapps.sjsu.edu/blogs/wp/vcara/) documents upcoming and past events. In addition, a Twitter account (Figure 5) (@Mnemosyne_SL) was created and is used, along with the blog, to announce upcoming events.

One of the biggest issues for documenting or preserving course elements is that of document spread. For example, we found that each student would create their machinima and upload it to their own social website, such as YouTube, and then provide the link, rather than the video itself in fulfillment of the assignment. This means that student videos were spread across the Internet with no control over preservation. A link can disappear at any time. Another issue is that of preserving objects created in-world. The instructor, mentors, and students all built numerous objects during the course of the semester which live in the creator's (avatar) inventory. Depending on permissions, the student may "give" or "sell" an object, but if not, once the owner picks up the object it is lost to anyone else who might want to view it or reuse it. At the end of each semester the entire build is taken down to make room for the next course to be taught. Additionally, sites like Spruz, offer free accounts, but these are so limited that an instructor would need to pay a monthly fee for an account that supports a typical class. This means that after the semester is over, the instructor account must be shut down if they do not want to continue to pay for class space.

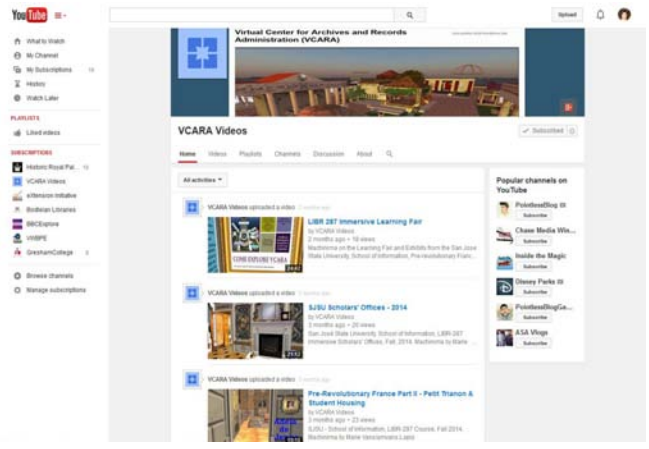

Figure 4. Virtual Center for Archiving and Records Management (VCARA) YouTube Channel.

Figure 5. Right: Pinterest Virtual Worlds Board for SJSU School of Information, Left: Twitter page

Figure 6. Virtual Center for Archives and Records Administration Blog page

Archiving Virtual Worlds Educational Data

In order to address the above issues, we have developed documentation & preservation scaffold consisting of a school-sponsored blog, centralized social media (YouTube, Pinterest, Word Press, twitter) and server storage space. The storage currently consists of frequently backed-up space available on one of the school's servers; however, we are currently looking at Amazon Glacier services, Google Nearline services, and Preservica to archive all of the data. The blog is used to document events and to provide links to machinima. To manage student-created machinima, students send the instructor the media file rather than a link. These videos are then uploaded to the central YouTube account. To document objects, students may give copies of objects for which they have permission to the instructor. These objects can be saved in a format suitable for storage and later re-imported. For objects the students are not able to transfer, a video can be made to document their existence and later stored with the rest of the videos documenting the course. Most events are captured in real-time and uploaded to the same social media account. Finally, a series of machinima are created specifically to capture the entire build (Tudor Village, Renaissance Italy, and Pre-revolutionary France).

Currently, YouTube does not have a limit on the number of videos that can be uploaded. The size limit is a generous 128GB, and the maximum duration is 11 hours.. This makes it ideal for public access to the videos and allows for others to see in more detail how such a course was delivered. However, there is no guarantee that this will not change in the future. Additionally, server space on the school's computer is limited, and the amount of data documenting the courses is rapidly increasing due to the size of video files. The same is true for the WordPress blog, the Pinterest account, & the Spruz class sites. In order to address the possible loss of data, we have created a process for archiving the data at the end of each semester:

1. All videos are saved in MP4 format and uploaded to the storage server along with all raw footage in the original capture format.
2. Images from the Spruz site and Pinterest site are uploaded to the storage server. They are culled to remove blurry or otherwise bad images as well as duplicates.
3. Pages from the Spruz site including text and images are saved as HTML files and uploaded to the appropriate semester directory on the storage server.
4. Pages from the WordPress blog are also saved as HTML files and uploaded to the storage server.

As previously mentioned, the amount of data, in gigabytes, has increased to the point where it is no longer feasible to store everything on the school server. We are currently looking at three options: 1. implementing a document and media archiving strategy utilizing the Amazon Glacier storage services from Amazon Web Services (AWS: aws.amazon.com), 2. using Preservica (preservica.com), which is an off-the-shelf digital preservation system built on top of Amazon Glacier storage services, and 3) contracting for Google Cloud Storage Nearline.

Amazon Glacier

The cost of service for Glacier is $0.01 per gigabyte of data stored per month. For a terabyte of data, this runs to about $10.00 per month. The exact cost depends on the number of times the data is accessed (uploaded or downloaded) and the number of bytes. This is extremely cost effective for our purposes as we will mostly be uploading data only a few times per year and accessed very infrequently. As the amount of data increases, the costs scale to the amount used, making it easier to plan for budget purposes. Data that is uploaded is transferred using SSL that automatically encrypts the data for security, although this is less of an issue for us. Importantly, the data is redundantly backed up on more than one system. The drawback of going with a cloud service like Glacier is that any archival standards and procedures would have to be developed locally. However, with a policy in place based on the 4 steps outlined above for each type of file and file structure, data capture, and uploading of data at the close of each semester, student volunteers would only need to be trained to use the Glacier interface. A second drawback is the retrieval time, which could be as long as several hours.

Preservica

Perservica is an archiving system that supports OAIS-compliant workflows. It is built on top of Amazon Glacier storage services, so automatically inherits the security and redundancy provided with it. Preservica addresses many issues inherent in the archiving of digital assets including support for migration of obsolete media and file formats, user-friendly data management, and automated data ingestion as well as a customizable interface

for use by internal and external users. Because this system has an archival process already built-in, the cost (in time) for setting up a process from scratch is greatly reduced. However, the cost in dollars is the biggest disadvantage to this solution. The most cost effective product is $3,950.00 per year for 100 gigabyte of data storage. For 250 gigabytes, the cost is $6,950. For us, the lion's share of data consists of video footage. Currently, our videos alone from three semesters consume 332 gigabytes. While compression could be applied to these files, in general, the 250 gigabyte limit for the more costly Preservica plan, seems small for us.

Google Cloud Storage Nearline

Google recently announced Google Cloud Storage Nearline, which matches Amazon Glacier's cost of service of $0.01 per gigabyte of data storage per month. But unlike Amazon Glacier's slow response time (several hours), Nearline promises to respond within 3 seconds. Although the 3 seconds is slow for web content, it is more than adequate for our purposes: long-term storage of digital objects infrequently accessed. Nearline offers redundant storage at multiple physical locations, and it is fully integrated with other Google Cloud Storage services.

Conclusions

Social media has become both popular and easy to use over the past few years, making it possible for instructors to conduct distant learning courses using technology with which prospective students are already familiar and at minimal cost. The problem is that course material and student-generated data become spread across the Internet and not very stable. We have shown how it becomes feasible to both use these social networking systems for learning and to contain the data for archiving. The price of cloud storage services is economical and its use almost effortless, removing the need to maintain computer equipment and manage updates. The procedure for archiving data at the end of the semester is also straightforward and could easily be automated using web automation software.

References

[1] Lowood, H. "Memento Mundi." Digital Media: Technological and Social Challenges of the Interactive World (2011a): 1.

[2] McDonough, J., Olendorf, R., Kirschenbaum, M., Kraus, K., Reside, D., Donahue, R., Phelps, A., Egert, C., Lowood, H., & Rojo, S. (2010). Preserving virtual worlds final report. University of Illinois. Retrieved March 26, 2015, from http://hdl.handle.net/2142/17097

[3] McDonough, J. P. "'Knee-Deep in the Data': Practical Problems in Applying the OAIS Reference Model to the Preservation of Computer Games." System Science (HICSS), 2012 45th Hawaii International Conference on. IEEE, 2012.

[4] Winget , M.A. & Murray, C. (2008). Collecting and preserving videogames and their related materials: A review of current practice, game-related archives and research projects. Cornell University Library. Retrieved March 26, 2015 from: http://arxiv.org/abs/0811.3137v1.

[5] McDonough, J. P. "Packaging videogames for long-term preservation: Integrating FRBR and the OAIS reference model." Journal of the American Society for Information Science and Technology 62.1 (2011): 171-184.

[6] Smith, M. "Curating architectural 3D CAD models." International Journal of Digital Curation 4.1 (2009): 98-106.

[7] Bruner, J.S. Toward a Theory of Instruction. Random House, New York, NY, (1966).

[8] Bandura, A. Social Foundations of thought and action: A social cognitive theory. Prentice Hall, Englewood Cliffs, NJ, (1986).

[9] Vygotsky, L.S. and Cole, M. Mind in Society: The development of higher psychological processes. Harvard University Press, Cambridge, MA, (1978).

[10] Wenger, E. McDermott, R. and Snyder, W.M. Cultivating Communities of Practice. Harvard Business School Publishing, Boston, MA, (2002).

[11] Theimer, K. "What is the Meaning of Archives 2.0?." American Archivist 74.1 (2011): 58-68.

[12] Zarro, M., and Hall, C.. "Pinterest: Social collecting for# linking# using# sharing." Proceedings of the 12th ACM/IEEE-CS joint conference on Digital Libraries. ACM, 2012.

[13] Jensen, B. "Instagram as Cultural Heritage: User Participation, Historical Documentation, and Curating in Museums and Archives through Social Media." 2013 Digital Heritage International Congress.

[14] Library of Congress. Prints and Photographs Division, et al. "For the common good: The Library of Congress Flickr pilot project." Library of Congress, Prints and Photographs Division, 2008.

[15] Lowood, H. "Perfect Capture: Three Takes on Replay, Machinima, and the History of Virtual Worlds. "Journal of visual culture, 10(1), 2011b, pp. 113-124. [1] John Doe, Recent Progress in Digital Halftoning II (IS&T, Springfield, VA, 1999) pg. 173.

Author Biography

Patricia Franks is an Associate Professor and coordinator of the Master of Archives and Records Administration (MARA) degree program in the School of Information at San José State University. Her interests lie in archives, records management, and emerging technologies. She has a Ph.D. in Organization and Management from Capella University.

Marie Vans is currently a Research Scientist with Hewlett-Packard Labs in Fort Collins, Colorado. Her main interests are security printing and document analytics. She is also a Masters Student at San José State University in the School of Information. She has a Ph.D. in computer science from Colorado State University.