
This paper describes an interview and observation-based study on the practices of communities of game players in preserving digital games using disk images and hardware emulators. An emulation-based preservation strategy was built based on the findings from this study.

Five hobbyist digital game preservers were interviewed via online text-based chat clients, two online forum-based discussions were conducted, and seven forums and eight websites dedicated to the practice of game preservation and the distribution of disk images or ROMs were read and analyzed. The game players were found to adhere to high standards of authenticity and data integrity and demonstrated a concern that not enough was being done to preserve digital games. The resulting strategy of emulation-based digital preservation derived from their practices could be adapted to the need of a heritage institution. However, more research is necessary before a complete strategy with an appropriate metadata structure adapted to the specific needs of digital games can be developed.

Headings:
Digital preservation -- Digital forensics
Digital preservation -- Video games
Internet research -- Interviews
Virtual communities -- Video game players
BUILDING A STRATEGY FOR DIGITAL GAME PRESERVATION BASED ON THE PRACTICES OF COMMUNITIES OF PLAYERS

by
Cherie Heiberg

A Master’s paper submitted to the faculty of the School of Information and Library Science of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Master of Science in Library Science.

Chapel Hill, North Carolina
November, 2013

Approved by:

________________________
Advisor
Introduction: On the Need for Preservation

“Emulators for early game consoles are generally available as freeware. These exist at a multitude of web sites, created, maintained and contributed to by amateur enthusiasts. The expertise here and the profusion of systems and software demonstrates what can be done in the area of emulation” (Ross & Gow, 1999).

Digital games have risen in both popularity and importance in the last decade. The term “digital games” encompasses console-based video games such as the Call of Duty series, computer or games such as the Oregon Trail playable on Apple IIe, arcade games like Pac-Man, mobile device-based games such as Angry Birds, and online games such as World of Warcraft and Candy Crush Saga. In spite of a lingering reputation of frivolity, these games are now on equal popular footing with movies and television. In 2010, the worldwide video game industry generated $25 billion in revenue (Entertainment Software Association, 2012b), and the Motion Picture Association of America reported $31.8 billion in worldwide box office sales for all films regardless of distributor. *Halo 3*, the bestselling digital game of 2007, grossed more revenue on its first day on the market than the top grossing film of 2007, *Spider-Man 3*, gained in its entire opening weekend (Barwick, Delaney, & Muir, 2008). Seventy-two percent of American households play digital games, and 82% of these gamers are 18 years of age or older (Entertainment Software Association, 2012a). Digital games are a part of our shared culture, and must be approached with as much attention towards preservation as other important cultural objects.
There has been no shortage of voices calling for the preservation of these games before they are lost (Anderson, Delve, & Pinchbeck, 2010; Barwick, Delaney, & Muir, 2008; Barton, 2005; Lowood, 2004; Newman, 2009; Newman, 2013). Those in popular media have also noticed that something is amiss: Super Mario Bros., a cartridge-based game released for the Nintendo Entertainment System (NES) that is largely credited with saving the languishing American video game console market, has no verifiable release date in the United States (Cifaldi, 2012; Risen, 2010). This is a game that has sold over 40 million copies worldwide since it was released (Guinness World Records, 2005). It is clear that something needs to be done.

It is only in the past few years that a concentrated effort to preserve digital games has emerged (Anderson, Delve, & Pinchbeck, 2010; Lee, Tennis, & Clarke, 2012; Lowood et al., 2009; McDonough et al., 2010; Murphy, 2013; Project KEEP, 2012). There is a widespread perception among preservationists that the issue of software preservation is too complex to be dealt with (Swalwell, 2009). This mindset is a major impediment to preservation work. It is also surrounded with a thorny sea of intellectual property issues that favor current rights owners over the interests of future generations (Barwick, Dearnley, & Muir, 2011). Additionally, digital games are so vastly heterogeneous - and sometimes novel - that few preservation standards exist as of yet (Barwick, Dearnley, & Muir, 2011; Lee, Tennis, & Clarke, 2013).

Though the effort to preserve digital games continues to build in professional preservation circles, communities of game players have been preserving digital games for decades (Lowood et al., 2009; Murphy, 2013; Newman, 2013). The online emulation
community is the largest source of historical game information and preservation that exists today (Goodling & Terras, 2008; Murphy, 2013; Newman, 2013). Members are both preservers and users of digital games. This classification represents the two most popular approaches in video game preservation: preserver-focused strategies (McDonough, Kirschenbaum, Reside, Fraistat, & Jerz, 2010) and user-focused strategies (Lee, Tennis, Clarke, & Carpenter, 2013). The individuals preserve digital games not only because of their passion for the genre, but due to the lack of interest from both the video game industry and preservation organizations (Barwick, Dearnley, & Muir, 2011; Risen, 2010). They preserve copies of digital games that were made for obsolete hardware environments, and, if needed, they recreate those hardware environments with hardware emulators. Many gamers tend to preserve not only the code of game itself, but the context of the game, such as scans of the box art, the instruction book, and screenshots of the game during play (Lowood, 2004; Goodling & Terras, 2008). They create metadata meant to keep the game accessible. They share information on the games that they are preserving with the larger digital gaming community in order to create more authoritative sources of information when there are none to be found. In doing all of this, they preserve the game, some of the context around it, describe it, and make it available for people of today to experience. They are, in short, behaving like archivists.

Thus far, however, there has been no serious work done towards the development of a preservation strategy based on the practices of communities of players. The practices of these communities of players obey the archival principles of
context and authenticity and demonstrate considerable technical skill. By studying their practices, this paper aims to develop a strategy for the preservation of digital games based on emulation of hardware and the preservation of original code. First there will be a review of the literature on current preservation strategies of digital games. Second, the research methodology used in this study will be described. Third, a description of current preservation strategies used by online communities of gamers, collected through a series of interviews and through observation, will be presented, and a proposed strategy for emulation-based digital preservation based on these practices will be produced. Lastly, there will be a discussion of the directions in which this strategy could further be developed, and the potential impact of communities of gamers on the state of game preservation as a whole.

**Part 1: A Review of the Literature**

**Existing Methods of Preservation**

In recent years, the professional digital preservation community has begun to develop various methods of preserving digital games. Technology preservation, or the storage of games and their components, is the most direct and common way to preserve digital games at this time (Bridge & Pruyn, 2008; Guttenbrunner, Becker & Rauber, 2010). Generally, a physical copy of a game is preserved along with a device that can be used to play the game. To preserve a copy of Pokemon Pearl, a Nintendo DS must also be preserved or the game will be inaccessible. This is not a long-term solution. Original hardware may fail, software may become damaged, and as more console-based games
are created, more space must be made available to accommodate them. It is in the best
ingests of a company based on hardware and software sales to focus on making new
technology that will replace the old rather than making something that will stand the
test of time (Goodling & Terras, 2008). Digital objects, especially commercial digital
games, are not often built with longevity in mind.

The most common preservation strategy for a digital object is migration (Granger, 2000). Migration is a strategy that takes digital objects and converts them to new
formats so that they do not become obsolete as technology progresses (Winget &
Murray, 2008). It involves the periodic transfer of data from an obsolete computing
platform, medium, or format to current platforms, media, or formats. Though it is
logically possible to indefinitely migrate complex digital objects from one platform to the
next, it is not possible to do this in reality. It is an extremely labor intensive process, and
requires near constant interference on the part of the archivist (Rothenberg, 1999).
There too many different types of digital objects dependent on too many different
software or hardware environments for this to ever continue smoothly (Anderson,
Delve, & Pinchback, 2010). There are other problems as well. The pacing of a digital
game, for example, can be affected by the improved processor speeds when it is
migrated to a new machine. This can create an entirely different gaming experience from
how it was originally intended, which in turn affects the authenticity of the object
(Winget & Murray, 2008; Hedstrom, Lee, Olson, & Lampe, 2006). Migration cannot be
seen as the best solution to the problem of digital preservation.

Emulation one of the most advocated solutions to this problem. Emulation,
which is similar to but distinct from simulation, is used to recreate the actions of the
original hardware in an accessible form. (Barwick, Dearnley, & Muir, 2011). Simulations
recreate the environments of both the hardware and software, whereas emulations
focus on recreating the design of the system or on creating an environment in which the
original software may run, on the system or the software level (Ross & Gow, 1999;
Winget & Murray, 2008). A full simulation of the Nintendo 64 game *Banjo-Kazooie*, for
example, would be a piece of software designed to run in a modern hardware
environment that recreates the original look and feel of the game as closely as possible.
An emulation of the same game would re-create the original Nintendo 64 environment
in order to provide a way to run the mostly-unmodified code of the game. In this way,
the original code of the game is preserved as a complete representation of the game
itself (Murphy, 2013).

“Emulators,” writes Murphy, “Are clever cheats that allow game code to be
accessed after its platform has broken down” (p. 47, 2009). By recreating the hardware
environment, emulators help to preserve the original digital game’s authenticity.
Emulation at the hardware level is the ideal solution for console games, the code of
which is not only wildly complicated (Anderson, Delve, & Pinchbeck, 2010), but often
protected by the rights holders. When it is properly used, emulation does not affect the
original code of the digital object that needs to be preserved, which protects it from
unintentional modification, and thus protects the object’s authenticity (Rothenberg,
1999).

It is the code of the digital game itself that must be preserved. Though it is
important to preserve the physical artifact from which it came, the code of a digital game can be considered the artifact itself rather than the physical platform of the game, whether it be disk, mobile phone, or cartridge (Murphy, 2013). This is due to the way one interacts with digital games as opposed to other mediums. The interactivity aspect of the game is akin to a performance (Bridge & Pruyn, 2009). A DVD that contains footage of a dance performance from the early 1920s cannot be considered the authentic artifact; rather, the footage itself is the artifact, because the footage can only be authentically experienced when it is viewed. Digital games must be experienced the same way. In the event that the original technology fails, the code within, if it has been properly preserved, serves as the artifact (Murphy, 2013). Authenticity and access are both maintained.

A method that remains largely unexplored save in the realms of digital forensics is the creation of disk images. Disk images are sector-by-sector copies of digital content from a physical storage medium (Woods, Lee, & Garfinkel, 2011). Stored disk images of this code can provide proof of file integrity and chain of custody, can serve as baseline comparisons when evaluating preservation actions, can provide backups in case of unexpected data loss, and can be shared between institutions (Woods, Lee, & Garfinkel, 2011). Disk images are also read-only; unalterable snapshots of the system within the physical device. Because of this, one can be assured that access to the image will not alter or damage the file. This enforces the legitimacy of code as artifact.

When this original code is combined with hardware emulation, the original code of the game is not affected. The use of a hardware emulator, commonly referred to
simply as an “emulator,” creates a single, multi-use environment that can be used to access multiple games created for the same original system, operating system, or computing platform, which in turn creates less work for the preserver (Guttenbrunner, Becker & Rauber, 2010). Using one emulator for hundreds of preserved games creates far less work for the preserver. Additionally, using emulation instead of full simulation is both less technologically challenging and puts less demand on the hardware of current operating systems (Barwick, Dearnley, & Muir, 2011; Ross & Gow, 1999; Winget & Murray, 2008). When the hardware environment for a digital game fails, as it inevitably will due to the custom parts that are necessary for it to run, lack of support from the company that created it, or simple technological deterioration (Barwick, Dearnley, & Muir, 2011; Guttenbrunner, Becker & Rauber, 2010), the emulator and the code of the game will remain.

However, the code of a digital game must be preserved with knowledge of the game’s original platform, or it runs the risk of violating the game’s authenticity by removing its context (Murphy, 2013). “To preserve an Atari 2600, do you need a piece of shag carpet?” asks Matt Kirschenbaum, collaborator in the Preserving Virtual Worlds project (Risen, 2010). When creating an emulation of a digital game, should one also preserve the box art and the instruction manual as relevant ephemera? What if the game is born digital? Should a snapshot of the online community it thrived in be somehow preserved as well? Can it be preserved? For console games, there is also the question of the controlling device, or “controller,” to consider. Can a digital game truly be preserved as close to its original context as possible without the use of the controller
that originally went with it? This question has been covered sparsely in the literature (Guttenbrunner, Becker & Rauber, 2010; Hedstrom, Lee, Olson, & Lampe, 2006; Murphy, 2013). The Multiple Arcade Machine Emulator (MAME) project approaches authenticity as attention to reproducing the original game as close to the original as possible, complete with bugs and glitches (Murphy, 2013). This can cause problems on modern PCs, as the technical intricacies of emulation are often quite technologically taxing (Lee, Olson, & Clifford, 2006; Murphy, 2013). As gaming technology continues to evolve, more and more questions will be raised about how to fully preserve them without suffering a huge loss of context or a risk of deterioration of the game thanks to the limits fragile commercial technology.

**Current Metadata Strategies**

The literature points to a clear need for improved metadata standards when documenting digital games. Digital games are beginning to be preserved very late in their existence. This means that the metadata standards used to catalog them must either be updated or created completely from scratch (Anderson, Delve, & Pinchback, 2010; Lee, Tennis, & Clarke, 2013; McDonough, 2012). A preserved digital object without the necessary metadata to access it is just as useless as a cassette tape without a tape player and an instruction book to learn how to use it. If nothing is done to address this problem, even if digital games and their hardware components are preserved, they will be unable to be accessed, and thus lost to history (Barwick, Delaney, & Muir, 2008).

Many preservation efforts have derived their methods of organization from two
sources: the field of knowledge organization, and commercial digital game companies found on the internet (Lee, Tennis, & Clarke, 2012). These sources are not without their problems. Knowledge organization is largely focused on the management of analog materials such as books or simple digital objects such as electronic articles. They lack a controlled vocabulary that applies to game genres, reliable sources of information, unclear region and language information, and a common inability to distinguish between publisher and developer (McDonough, Kirschenbaum, Reside, Fraistat, & Jerz, 2010). Commercial systems have no standards across companies, and are often contradictory in their metadata terminology or lacking in credible sources of information (Lee, Tennis, Clarke & Carpenter 2013). These models of organization cannot be easily applied to the specific needs of digital games, which include such issues as versioning conflict, lack of a compatible hardware environment, missing instruction manuals, and a dearth of authoritative sources of information.

One of the more interesting efforts put forth to build a viable metadata structure is by Lee, Tennis, Clarke, and Carpenter for the Seattle Interactive Media Museum (2013). Their focus was to create access points for digital games from a user’s standpoint rather than a preservation standpoint. To that end, the members of the study created personas, or archetypes based on the needs, behaviors, and goals of particular group of users, and measured their perceived needs against existing metadata elements based on aggregated data from commercial, review, and hobby websites. Using personas instead of interacting with actual gamers may have limited the accuracy of their chosen metadata elements, due to possible limited understanding of the actual needs of their
 personas and a risk that their needs would be too broad. They plan in the future to conduct full-scale interviews with users to fully assess their needs.

**Copyright**

The largest obstacle archivists have faced in learning from communities of players to build preservation strategies is the fear of being associated with software piracy. Disks and disk images or commercial products such as games are subject to copyright law. If a disk image is created and distributed without the permission of the rights holders, the creator of the image would be subject to penalties under the jurisdiction of the law (Woods, Lee, & Garfinkel, 2011). Those who create and distribute images or repackaged versions of digital games online only continue to exist due to oversight from the companies that produce the games (Barton, 2005; Gooding & Terras, 2008). Large companies like Nintendo, Sony, and Microsoft are generally focused on the profitability of their current technology over the profitability of technology that they no longer support. However, they would be well within their rights to enforce their copyright at any given time, which happened with a company called Cloanto, that bought the executive license to Amiga operating system ROMs and sent cease and desist orders to anyone on the internet who had distributed them (Barton, 2005). An archive that exclusively used emulators built by amateurs would have to remain closed until such a time as the copyright on the game expired or permission was attained from the company that produced the game to open it to the public. The MAME project remains a closed archive for this reason (Murphy 2009). As long as this grey area exists in which a
company may or may not pursue legal action against people who violate their property rights, it will be difficult for archival or heritage institutions to commit time and resources to developing an emulation-based preservation strategy.

To make matters worse, it takes a considerable amount of effort to track down and determine the original rights holders for games that have been produced as little as a decade ago. Game companies are not permanent institutions. They may have gone bankrupt and vanished, or those who held the original rights to the game may have moved on from producing digital games at all (Barton, 2005; Gooding & Terras, 2008). Additionally, technological measures to enforce copyright protection have been in place since the earliest days of the digital game medium (Newman, 2013). Without information from the company that produced the game on how to crack the copy protection, one would have to dedicated considerable time to experimenting with various copy-protection breaking techniques, which are time-consuming, faulty, and illegal. This is in addition to the issue that “cracked” or “hacked” software can present problems upon emulation, such as the loss of essential parts of code (Lowood, 2004). If permission cannot be obtained, and documentation about the game cannot be found, it is difficult at best to preserve a game as anything but a piece of decaying technology.

There is, however, some hope. In 2003, the U. S. Copyright Office ruled that video and computer games that are considered obsolete may be digitally preserved without violating copyright law (2011; Moore, 2003). This law was put in place essentially to protect magnetic media such as floppy disks, which have a lifespan of between 10 and 30 years, from technological oblivion (Internet Archive, n.d.; Ross & Gow, 1999). This
ruling was last made in 2010, and is renegotiated roughly once every three years. It has been renewed twice since 2003. Archives and other heritage institutions such as the Internet Archive are able to utilize this exception to copyright law to justify their efforts to preserve digital games (Moore, 2003; Internet Archive, n.d.). Emulation has been largely ignored because of an overly-careful approach when it comes to the violation of copyright (Murphy, 2013). If this exception exists, then archivists should take advantage of it. There are already communities of people online who have been creating emulators and disk images since the early days of gaming. Newman's call in his paper on piracy as preservation for greater collaboration between preservationists, game professionals, and communities of players is one that should be answered (2013).

Part 2: Methodology

For roughly one month, I sought out people online who preserve digital games and use hardware emulators to play them. I recruited people to be interviewed by posting one of two versions of a recruitment letter on the general talk section of online discussion forums attached to websites that are dedicated to the preservation of digital games. I also directly emailed several people who are well-known in the preservation community (See Appendices 1 and 2). I was able to identify these people with the help of players who expressed interest in my project but who did not create hardware emulators or preserve game code. I also interacted in forum-based discussions with people who were unwilling to be directly interviewed, but who were interested in discussing the topic of preservation with full awareness that I was conducting research.
Over the course of these months, I interviewed five people over eighteen years of age from Germany, Canada, and the United States, and participated in two forum discussions.

My interview questions focused on several aspects of game preservation (See Appendix 3). The first three questions were designed to evaluate the level of experience of the player being interviewed and to assess the motivations behind their work. The next three questions were asked to determine how the person being interviewed preserved games or how the person made hardware emulators. The next three asked about context and the methods by which players procured information about the games that they preserved. The last four questions asked about the effect of copyright on preservation work and the gamers’ opinions on the state of video game preservation.

I also carried out further research and observation by reading through the archived threads in forums and studying websites dedicated to game preservation. These forums were Abandonia, VOGONS, LostLevels, LemonAmiga, Zophar’s Domain, The Old School Emulation Center (TOSEC), and the Good Old Days. The topics I read covered subjects related to the creation of hardware emulators, the creation of disk images and ROMs, discovering information about existing games, discussions on copyright, and ephemera collected about games. The websites that did not have forums but covered the creation of emulators or disk images that I researched were Tiger Emulators, Kult Cover Disks, Demu, RGB Classic Games, Dosgames.com, DOS Games Archive, and Redump.org.
Part 3: Findings

Assumptions

I did not begin this project without bias or without assumptions. I am a game player myself. The first digital game that I can remember playing was King’s Quest 1: Quest for the Crown on my father’s work computer some time in 1992. Digital games have been a chosen method of leisure for me since then. I have a vested personal interest in preserving digital games. I have spent time playing digital copies of classic games that I procured through Abandonware sites. Because I have been exposed to a small amount of these Abandonware websites, my assumption was that there would be a unified method of preservation utilized by everyone who made it their business to preserve digital games. In turn, I assumed that most, if not all preservers of these games would hold a blasé attitude about copyright infringement, not because of a personal interest in violating copyright laws, but because if they did not preserve these games, who would? Because of these assumptions, I included in my interview questions inquiries about copyright law, preservation practices, and the methods by which preservers of games learned their skills.

I also assumed that I would have trouble recruiting people to be interviewed. I saw the online game preservation community as insular and distrustful of outsiders for a number of reasons. One, that they are keenly aware that what they do could be considered illegal, and that they have no wish to make contact with someone who could put them into jeopardy with copyright law. Two, that although I am a game player, I do not have an active online presence in any online community that deals specifically in
game preservation. Three, that because I am not an insider, I was likely to use outsider
language about game preservation that would lead to misunderstandings and potential conflict. With these assumptions in mind, I attempted to keep the language of my recruitment letter as neutral as possible, using terminology such as “Abandonware” or only when I could find no alternative term to use. Abandonware is a term that has legal implications that many gamers are not comfortable with. I used it in one forum and was met with a negative reaction. I therefore refrained from using it in future interactions, instead using the term “digital preservation of games” when necessary.

Communities of Players

The online digital game preservation community is not a single, unified entity. It is composed of a wide variety of people from multiple locations in the world, all of whom have their own particular focus in preserving games and the ephemera associated with them. Some groups create modified versions of games that can be played as executable files, some strictly preserve original distributions of games, some collect original copies of games and create image files, some create hardware emulators for modern computing environments and mobile devices, some collect and preserve hardware emulators, etc. One person I encountered in a forum described his particular interest as preserving game manuals, which are instructional books that often are packaged with digital games. They often aid the player in figuring out basic controls and include such pertinent information as the date of publication, the producers of the game, and the company that published it. He preferred at all times to handle the original
game manual as little as possible in order to prevent damaging them accidentally with tears and creases. When he needed a manual, he would seek out or create a digital copy, print it, and bind it, similar to the archival practice of creating “access copies” of fragile multimedia records (Gustman et al., 2002). The varying factions of preservationists are proud of what they do, and do not want to be confused with preservationists who engage in methods or focus on areas other than the ones with which they are particularly concerned.

Abandonware serves as a good illustration of this. Abandonware, which has received coverage in the literature (Barton, 2005; Barwick, Dearnley, & Muir, 2008; Lowood, 2004; Lowood et al., 2009), refers to a specific type of game, and cannot be applied across all games. Abandonwarez are games that have been essentially “abandoned” by their creators. They are longer commercially available, or they were made by companies that no longer exist, or it is uncertain who owns the copyright at all. Therefore, they are generally seen as existing in the grey area outside copyright law. A game that is owned by a company that no longer exists cannot be easily protected, because there is no one left to enforce the copyright (Barwick, Dearnley, & Muir, 2008). This does not mean that the copyright cannot be enforced, but it does mean that it is extremely unlikely. Abandonwarez, therefore, are often treated as “ownerless” and copied, distributed, and sometimes edited by people interested in playing them.

Other game preservers avoid using games under copyright. For example, the creators of DOSBox, a hardware emulator built to recreate the environment of an Intel x86 PC, work closely with the users of their emulator to ensure that the games played on
it are either original distributions or the unmodified images thereof if the original technology can no longer be accessed, as in the case of 5.25 inch floppy disks. In their forum, the word “Abandonware” is strictly verboten.

The difference between these two approaches to digital game preservation stem from ideological differences similar to those of archivists. Supporters of Abandonware and all that it entails fall firmly on the side of access over authenticity. This is not to say that authenticity is not important to users of Abandonware. Some creators of Abandonware are adamant about creating only image files and not altering the original code of the game. Indeed, there are many similarities between those who utilize Abandonware in their preservation efforts and those who only use unmodified, original distributions. Both show concern for in-tact code and for graphics and sound that reproduce the look and feel of the game in its original environment. The largest difference is that users of Abandonware report that their work in preserving games cannot afford to wait. One preserver who works from Germany stated that while he did worry about copyright law, that if he and his colleagues did not preserve these games, then who would?

Those who do not support Abandonware, while they too see themselves as doing important preservation-related work, are strongly opposed to violating copyright law. They are also wary of accepting any type of Abandonware as authentic, as the code may have been modified from its original version. This is generally done to update it or to cut down its file size, among other possible reasons. Without the original version to serve as comparison or accompanying metadata, there is no way of knowing whether a code has
been modified, whether it is a demo, the full game, an incomplete image file, or whether parts of the game have been cut. There is also no way of knowing whether downloading it in the first place is a violation of an international copyright law. Authenticity and integrity, to them, are key tenets of preservation. Accessibility may suffer as games become lost while the search for their original versions continues, but file integrity is never violated, and the law is obeyed.

**Preservation Techniques**

The most commonly used way of preserving a digital game is to create an image of it. The creation of a disk image is a relatively simple procedure as opposed to ripping the code from its original location and rewriting it to adhere to current operating system standards, especially considering the relatively recent online proliferation of open-source digital forensics tools (Woods, Lee, & Garfinkel, 2011). This is done in a number of ways, depending on the source of the original game. DOSBox users commonly work from games that were originally distributed on 5.25 or 3.5 inch floppy disks. Disk images created from a 3.5 inch floppy can generally be accessed using an external floppy drive connected to a computer through USB. Modern motherboards still have the built-in capacity to communicate with 3.5 inch floppy drives. They cannot, however, read 5.25 inch floppies; they lack the necessary controller on the motherboard (John, 2008). One preserver, who works as an electrical engineer in his professional life, is able to do use his own expertise and training to get past these technological hurdles. However, while there are no official companies that produce devices that can be used to read 5.25 inch
floppies on modern computers, there are a few individuals who create controllers and sell them online (John, 2008). They may be purchased and used with software like the FTK Imager to create images if the disk has not deteriorated beyond recovery (Woods, Lee, & Garfinkel, 2011; Ross & Gow, 1999). Once the image is created, it can be preserved on an online server and distributed without fear of violating the integrity of the original code.

Copy protection is sometimes an issue. This becomes more difficult the more technology advances. With compact disks and digital distributions of modern games, copy-protection can often be built into the code of the game, or can be built into the architecture of the disk (Moore, 2003). There are methods of breaking copy protection, though contacting the original copyright holder and requesting permission is the ideal solution. One preserver, who specializes in creating images of CDs packaged with gaming magazines, found that he could break the copy-protection on some of these CDs by licking it off the disk. Through experimentation, he discovered that each disk he could not image had a transparent layer over the bottom that he could feel but could not see. He has not tried the technique with DVDs so far, as they are outside of his scope of interest, but suspects that it may work. Most other optical media must be imaged or ripped from the disk using specialized software that is available freely on the internet, developed by other game players or by people who perform software piracy.

Cartridge-based games present another problem: the bridge between a computer environment and a game cartridge built for a particular system such as the NES or Sega Genesis is a difficult one to cross. The files created from cartridge-based
games are known as ROM images, or just “ROMs,” because they were created from the read-only memory microchip present in the cartridge (Guttenbrunner, Becker & Rauber, 2010). Those who create ROMs from cartridges must use peripheral technology such as Retrode (Retrode.org, 2013) because a modern machine will not have a built-in port capable of reading a game cartridge. Copy protection also presents a problem. Various encryption methods were built into the ROM chips on cartridge games in order to prevent piracy. Some of these codes are very difficult to crack, or if cracked, could lead to damage to the ROM chip. The MAME team has managed to crack many encryption codes of arcade machines over the years (Murphy, 2013). Abandonware preservers have experience in breaking encryption when necessary. Players who engage in this kind of preservation work often learn from each other the best methods of breaking encryption, or from former employees of defunct game companies who have leaked information about their companies’ encryption methods to the web.

Born digital games, including ones distributed on the gaming networks such as Playstation Network and Microsoft Online, come with copy protection issues of their own. Born digital games or games that require an internet connection cannot be played once official support from their parent company ceases. Some publishers release patches near the end of the game’s lifecycle to ensure that the game will not vanish, but this is not a universal practice. Though there is a proliferation of them on the market today, because of their digital nature, when the company behind the games wishes to stop distributing them, they will be extraordinarily difficult to locate and preserve. Additionally, because the games are born digital, they do not come with any contextual
material such as a manual or a case. Copyright concerns keep many from going this route today, along with comfort derived from their proliferation level, but there is already a growing concern about this issue among online preservationists.

Creators of hardware emulators produce a vital half of the emulation equation. Creating an emulator is fantastically complicated when compared to creating a disk image of a digital game. Moreover, those who create hardware emulators are in arguably greater danger of violating copyright law than those who create disk images or who utilize Abandonware in their preservation work. Hardware emulators generally cannot be made without insider knowledge of some kind from the company that produced the hardware. This information can be found in company-specific documentation, which can be recovered from spec manuals leaked online by employees. The CD-i emulator, which recreates the Philips CD-i, was produced by a former employee of the company who is assumed to have had access to internal documentation. If documentation cannot be found, then hardware may be reverse engineered, which often results in the hardware’s destruction and an emulator with basic functionality.

Some open-source projects such as DOSBox have created hardware emulators through collaboration that are extremely close to the full functionality of the environment they are attempting to emulate, and which are in constant development as more aspects of the original environment are reverse engineered, recreated, tested, and added to the each new build of the emulator. Most importantly for game preservation, emulators provide stable environments in which to run disk images, which enables access to a game without compromising its authenticity.
Context and Ephemera

Game players who preserve games have a wide scope of interests outside of preserving the game itself. They collect ephemera such as gaming magazines, promotional posters and flyers, cardboard standees, Shareware and game demos, and other merchandize related to the game. Many “special” or “limited” editions of games, usually released in a limited quantity upon the game’s debut, come with ephemera such as cloth maps, limited download codes, jewelry, figurines, or other materials related to the story contained within the game. It is not uncommon to find threads in forums of someone seeking out “complete” copies of games that they enjoy, or to find an eBay listing of a complete limited edition of a game going up in price by the minute. One game player seeking out a complete copy of the rare adventure game, Cassandra Galleries, contacted the original developer of the game, Kutoka Interactive, who informed that they could offer him a copy of the game burnt from the original master disk for the cost of forty dollars, but this would not include any of the original contextual items. He declined their offer, already having an incomplete copy of the game. This collector’s attitude is very valuable for contextual purposes, and is largely understood by the gaming community to be an important activity.

The desire for “completeness” of a game is one reason why this attitude prevails, along with an understanding that a game preserved without as much contextual information as possible is in danger of not being fully understood. Scarcity of an item has been brought up as a reason to preserve it, as in the case of the gamer who preserves manuals of DOS games and creates copies of them so he does not damage his originals
during use. While there does not seem to be a prevalent attitude that preserving contextual information helps to recreate a snapshot of the time in which the game was originally played, contextual information has been used to provide a richer sense of understanding of a given digital game.

**Sources of Information**

There is currently a wealth of information available for games produced today. Games can commonly be found with their original packaging, with manuals and cover art intact, and the companies that have produced the game can be contacted if any information is lacking. For older games produced by extant companies, an online preservationist who has a game without its packaging or manual can often contact the company for additional information about the game, such as its year of production, the team behind it, its official name, or other information. That is, if the information is not already available online, through the game company’s official website (Lee, Tennis & Clarke, 2012). As we know from Cifaldi’s article, however, this does not always work (2012). Neither Nintendo nor Nintendo of America have records of Super Mario Bros. United States release date. At times like this, online preservationists have no recourse but one another.

For older games, authoritative sources of information can be few and far between. Sometimes the only pieces of information remaining about a game are found in the game itself. Over time, the box and manual were lost, or the company that owned the game folded and sold the copyright, or they folded and took the copyright, and any
authoritative source of information about the game, with them. Sometimes, even if the game code is available for exploration, it contains only hints as to the identification of the company that produced it. This challenge often causes game players to rely on each other to provide sources of information about the games that they wish to preserve.

The two main ways authoritative information about a game can be found when all other avenues are exhausted are turning to player-maintained databases and to other game players. Consulting one another for information about a game is a common practice. To an archivist, this may seem unorthodox, but consulting someone who has been playing games for more than ten years can yield excellent, accurate results. Answers can be confirmed by consulting other game players as well, leading to information that is as accurate as it can possibly be without direct information from the company that produced it. It can even be more accurate, as many game companies do not have record keeping systems and would not have had them in their early days. Even SquareEnix, a successful multi-national game corporation formed after a merger of Square Co., Ltd. and the Enix Corporation in 2003, did not preserve the game data of Kingdom Hearts, a popular game produced by the Square Co. in 2002 (Karmali, 2013, June 27). In contrast, game players have built online databases for quick consultation when they are in need of information about a particular game. While some databases are known to be poorly organized and badly maintained, others have good reputations among the players that use them, such as the Hall of Light (HOL Team, 2013) and Theodor Lauppert's Game Gallery (Lauppert, 2013).

The inconsistency of reliable information does create a problem for one who is
not already immersed in an online gaming community or who is not familiar with gaming culture. Game players are sometimes forced to use what librarians and archivists would not consider to be authoritative sources of information to discover and document aspects of the games that they preserve. There is no outright guarantee that the information obtained about a game was reliably procured. Indeed; while many online preservationists are concerned about the accuracy of their information, some are only concerned with accessing and playing the game, and they do not put effort into recording information unless it pertains to playing the game in the first place. With enough time and immersion one can eke out which databases are reliable and which ones are not.

**A Strategy for Emulation-Based Preservation**

Based on these practices of these communities of players, a strategy for emulation-based preservation of digital games emerges that can be adapted to the needs of heritage institutions, including archives and digital repositories.

A copy of the original game must be obtained. Ideally, the game will include its original packaging and, if applicable, the instruction manual that was packaged with it. Both items contain important publication and contextual information that are necessary for an accurate description of the game. While the title screen of the game itself can sometimes be used as a way to find this information, the game must sometimes be completed before the information is visible and revealed in the game’s final credits. If there is only an image file or a physical copy available to work with, more information
about the game can be discovered by contacting the copyright holder, the company that produced the game, other persons involved in the creation of the game as can be determined from the game itself, checking player-maintained databases, or seeking out help from communities of players online.

Once these materials are obtained, the format of the game must be identified in order to determine how it was originally accessed, whether it is still viable, what technology will need to be used in order to create an image of the game, and what hardware emulated will be used to access the game. If there is no contextual information available as described above, the format of the material can usually be visually identified by someone who has played digital games, or can be identified by checking online resources or consulting with online communities of game players.

There are two paths that can be taken at this point. If there is already an image of the game available online, it would be an ideal solution to collect the available image. This would save time and expense on the part of the institution. This should be done after it has been determined that the image in question is authentic, and after seeking out permission from the creator. If there is not an image available, if it is outside the scope of the institution to make such requests, or if it cannot be accurately determined whether or not the image is authentic, then an image must be created from the original game.

At this point an appropriate hardware emulator must be selected. Building a hardware emulator is a very complicated task. It should be completed with the aid of and permission of the company that produced the hardware. This may be an impossible
task for most collecting institutions. Because of the work of communities of game players, however, there is a wealth of hardware emulators that have already been produced available online. These emulators can be sought out, evaluated for their ability to faithfully recreate the original environment of the game, render the game once it is loaded, and played with as few bugs as possible (Guttenbrunner, Becker, & Rauber, 2010). Permission should be sought from the creators of the emulators to collect and utilize them. It is not recommended that current-generation consoles be emulated due to possible copyright issues that could emerge. Instead, they should be technologically preserved as long as is possible. As they become obsolete and official support for them vanishes as the copyright owners move on to more profitable ventures, a proper emulator can be selected and utilized.

Once the format is identified and a hardware emulator obtained, the method by which the image should be made can be determined. A computer capable of imaging must be used, software to enable the imaging process must be identified and obtained, and an imaging peripheral, if necessary, must be obtained or created. FTK Imager is a free program that can work on Windows and Linux machines, and the Disk Utility that is included with Macintosh operating systems (Woods, Lee, & Garfinkel, 2011). These programs can be used to create images of optical media and floppy disks. For the creation of ROM images, an appropriate peripheral technology must be obtained. The Retrode, a legal peripheral designed to create backup images of cartridge games, can be purchased for this purpose (Retrode.org, 2013). In the case of more obscure formats of games, the institution may need to seek out devices built by hobbyists, which could be
found by consulting with people involved with online game preservation work. If the
game in question is born digital, in the case of such games as Braid (Microsoft Online)
and LostWinds (WiiWare), it is best to preserve the code as-is, without any kinds of
modifications.

Metadata about the image of the game must be extracted and recorded both
with the preserved object and in a database. This mimics the use of websites to organize
information, and the practice of packaging a text-based “read me” file with some ROMs
or game images. The read-me file is not added to the image, but rather packaged with it
in a compressed zip file. Recording the metadata in two locations helps to ensure that it
is not accidentally lost to bit rot or accidental deletion. Basic identifying information
such as the name of the game, the year it was released, its system requirements, what
hardware emulator to use in order to access it, and the creators of the game including
producer and distribution company should be gathered from the game itself, its
packaging, or from the company that produced it. Legal information such as the status of
the disk image, the legal status of third-party requirements, and runtime rights need to
be documented as well (Anderson, Delve, & Pinchbeck, 2010). More interpretive
information such as the game’s genre should be applied based on a faceted schema,
such as the one being developed for the Seattle Interactive Media Museum (Lee, Tennis,
Clarke, & Carpenter, 2013).

Contextual information should be collected as well. This is important because as
time passes from the creation of the game, it becomes less and less clear how the game
was originally played, what impact the game had on popular culture, or why it matters
that it is to be preserved in the first place. Communities of players collect and produce a
great deal of ephemera out of love for the game. This can be an invaluable resource for a
person who wishes to build contextual information. Machinima, or short films made by
game players based on gaming technology, along with archived forum threads, can
provide good snapshots of community involvement of a game. Again, game players can
provide the most expertise in the context surrounding a game, and should be consulted
when available.

Once everything is collected according to the needs of the institution, there is
long-term preservation to consider. Communities of players ensure viability of the
images that they create by keeping local copies and utilizing personal or remote internet
hosting to allow others to download copies of the images. On an institutional level, this
can be interpreted as finding other institutions to host copies of disk images and the
metadata attached to them. This digital preservation strategy has been recommended
many times in the literature, specifically the LOCKSS system (Maniatis, Roussopoulos,
Giuli, Rosenthal, & Baker, 2005) and in the documentation strategies suggested by Helen
Samuels (1986). Digital forensics techniques could be applied to ensure the integrity and
authenticity of a disk image over a long period of time. Hexadecimal checksums should
be generated for each image and stored both with the image’s documentation and in a
separate file (Woods, Lee, & Garfinkel, 2011). If an image has succumbed to bit-rot, it
can be replaced with an identical image, or a new image can be generated from the
original game, if available.
Part 4: Conclusions and Directions for Further Development

There are several areas that were not adequately addressed in this paper that subsequent researchers could further develop. This paper did not investigate a metadata strategy, which could be developed by studying the organizational habits of communities of communities of game players. Game players are sensitive to the special description needs of digital games, and their understanding of this could help to create a deep and complete metadata system. Lee et al. have written that they plan to interview game players to further develop their metadata schema (2013), but it has not yet been done. I was unable to interview anyone who creates hardware emulators, and thus unable to bring their perspective into developing a preservation strategy based on emulation. The creators of emulators both develop and support the emulators that they create, and without their agreement to participate in institution-based preservation, an emulation-based strategy cannot be considered viable. It is also important to bring in the perspective of the companies that produce digital games. Obeying copyright is extremely important for heritage institutions. If an institution can gain permission from the copyright holder to create an image of the game they have produced, they can be assured that they are acting within the confines of the law. They could also possibly obtain valuable proprietary information about a game, such as information about copy protection and encryption.

There is currently a huge variety of games out there that must be preserved. From games on floppy disks, to cartridge games, games on optical media, games on SD cards, games for mobile phones, computer games that have huge worlds rooted in
distant servers, and more. The online preservation community has done a great deal of work in preserving what they can, but they are limited in their scope due to copyright restrictions, a lack of funding, and inability to access a dedicated server that could be counted on to keep their work preserved on a permanent basis. Virtual worlds, such as those tied to Massively Multiplayer Online Games like World of Warcraft, have not been a focus so far to communities of players. The worlds attached to the game along with the community functionality cannot be adequately preserved with disk images and emulators. The Preserving Virtual Worlds project has made strides towards answering the question of what to preserve when faced with games like this (McDonough et al., 2010), and Winget’s review of the literature reveals a movement in the professional field towards addressing this problem (2011). As technology evolves, and time passes between a game’s release and its obsolescence, the practices of communities of players to preserve these games evolve as well.

Without a more concentrated effort, it is likely that the professional world will always lag behind that of the hobbyist world. This is not due to a lack of interest. Professional efforts need funding to stay viable. Ones that have made strides in the field of emulation such as Project KEEP (2012) and the modular emulator Dioscuri (Koninklijke Bibliotheek, 2010) have concluded, and sit waiting to be further developed by someone who has the inclination and funding to do so. Preserving Virtual Worlds, which is now engaged in its second part, represents the most promising ongoing effort in digital preservation of games (McDonough et al., 2010), but it needs to go further and bring in the experiences of actual game players who have been involved in digital preservation
for more than a decade.

This paper has only scratched the surface of the wealth of expertise held by game players who preserve digital copies of games. If at all possible, it would be wise to bring on these game players as paid consultants for the development of digital preservation strategies of complex digital objects with hardware and software interdependencies, such as digital games and pieces of software. What these communities of players have accomplished is not easy. They developed their preservation techniques without getting paid and with a great deal of labor and time, knowingly risking punishment under the duress of copyright law. It is time that they were brought in for their expertise and paid for it.

The preservation work of communities of players is too prolific to be ignored. The knowledge and expertise that these groups have acquired are invaluable to future emulation-based preservation efforts (Barwick, Dearnley, & Muir, 2011). These groups should be working together with heritage institutions. The Digital Games Archive and the Internet Archive are both collecting commercial software with the permission of rights holders, but they are also mining the multitude of emulators made available by hobbyists (Lowood, 2004). However, in spite of the fact that archives can take advantage of the DMCA exception that grants them the ability to circumvent copyright law for preservation purposes, very little has been done in this area. The work of preserving complex digital objects such as emulated digital games will, in the end, fall to preservation professionals. By reviewing and documenting the practices of online communities of gamers, preservation professionals will make long strides towards
developing a “best practice” strategy of preserving not only digital games, but other emulated digital objects.
Bibliography


http://www.freesoftwaremagazine.com/articles/ liberating_games?page=0%2C0


http://www.gamasutra.com/view/feature/167392/sad_but_true_we_cant_prove_when_-php


Entertainment Software Association. (2012). Sales and genre data. 2013, from


Ponencia presentada en las actas de Fifth International Conference on Preservation of Digital Objects (pp. 48-55).

http://www.ign.com/articles/2013/06/27/original-kingdom-hearts-assets-lost

http://dioscuri.sourceforge.net/


Risen, C. (2010). Pac rat: The fight to preserve old video games from bit rot,


and Technology, 45 (1), 1-9.


Glossary of Terms

- **Abandonware** – Obsolete games produced by companies that no longer exist. Considered a “grey area” in copyright law, because it is not clear who should be contacted regarding the rights to copy and distribute the games in question.

- **Digital Games** – Games designed to be played on commercial consoles such as the Xbox 360 or the Super Nintendo, games played on mobile devices, and games played on computers or other stand-alone electronic devices.

- **Disk images** – Sector by sector copies of data that have been stored on a physical medium such as a hard drive, a floppy disk, a USB drive, a DC, or a DVD.

- **Emulation-based Preservation** – A preservation strategy for digital games that uses images of games and hardware emulators to recreate the game’s original environment without damaging the code.

- **Game manuals** – Small, bound books or papers that come packaged with games. Often contain instructions on how to play the game, information on who produced the game, and other contents related to gameplay.

- **Hardware emulators** – Sometimes called simply “emulators,” these programs are designed to emulate the properties of hardware or game consoles.

- **Open-source** – Generally speaking, this term refers to computer technology such as software or operating systems that are created and distributed without charge and with full access to the program’s source code.

- **Preservationist** – A paper-specific term referring to the communities of professional preservers of information, such as archivists both traditional and digital, digital librarians, and other curators of information.

- **ROMs** – Images of read-only memory chips found inside cartridge-based video games and arcade machines.

- **Shareware** – Complete games that are provided to users on a limited or trial basis for free or for a small fee which can only be experienced fully if the user pays a fee to the company that produced the game for a full licensed copy.
Appendix 1: Recruitment Letter 1

Hello! My name is Cherie Heiberg, and I’m a Masters student at the School of Information and Library Science at the University of North Carolina Chapel Hill. I’m writing my Masters paper on emulators. Specifically, I’m writing about how hobbyist creators of emulators and ROMs preserve video games. I’ll be comparing the work that hobbyists do to the work of professionals. The goal is to start building a professional model that can be used to preserve video games using emulators. A lot of the work will involve me observing what goes on in online communities that make emulators or deal in Abandonware. I’ll be looking at what kind of games get preserved, what kind of information/metadata gets attached to these games, how they’re distributed online, and how people get access to them. The most important part of the project, though, will be talking to the people who actually create ROMs and emulators.

That’s where you come in! I’m looking for people who would be willing to be interviewed about their process of creating ROMs and/or emulators. If you are over 18, I really want to talk to you! I won’t be collecting any information that could identify you, like your real name or your home city or your birthdate. You don’t even have to use your handle; you can pick out a completely new pseudonym and that’s good enough for me. I want to know why you do what you do and how you do it. If you’re interested, let me know by emailing me at heiberg@email.unc.edu or responding to this post with your chat handle and your preferred chat client, and we can set up an interview by email or by a chat client of your choice.

-Cherie Heiberg
UNC SILS 2013
Appendix 2: Recruitment letter 2

Hello! My name is Cherie Heiberg, and I’m a Masters student at the School of Information and Library Science at the University of North Carolina Chapel Hill. I’m writing my Masters paper on emulators. Specifically, I’m writing about how online creators and users of emulators preserve video games. I’ll be comparing the work that online creators do to the work of professionals. The goal is to start building a professional model that can be used to preserve video games using emulators. A lot of the work will involve me observing what goes on in online communities that make emulators or preserve emulated games. I’ll be looking at what kind of games get preserved, what kind of information/metadata gets attached to these games, how they’re distributed online, and how people get access to them. The most important part of the project, though, will be talking to the people who actually create emulators and preserve games.

That’s where you come in! I’m looking for people who would be willing to be interviewed about their process of creating ROMs and/or emulators, about the creation of disk images, or the preservation of original distributions of digital games. If you are over 18, I really want to talk to you! I won’t be collecting any information that could identify you, like your real name or your home city or your birthdate. You don’t even have to use your forum handle; you can pick out a completely new pseudonym and that’s good enough for me. I want to know why you do what you do and how you do it. If you’re interested, let me know by emailing me at [email]heiberg@email.unc.edu[/email] or responding to this post with your chat handle and your preferred chat client, and we can set up an interview by email or by a chat client of your choice.

I use skype, IRC, and gchat but I can set up an account for anything you like if you're interested.

Thank you for taking the time to read this!

-Cherie Heiberg
UNC SILS 2013
Appendix 3: Interview Questions

1. How long have you been playing video games?
2. How long have you been making emulators/ROMs/images?
3. Why do you make emulators/ROMs/images?
4. How did you learn to make emulators/ROMs/images?
5. What is your usual process of creating an emulator/ROM/image? Clarifying questions: How do you harvest the original code? What kind of equipment do you use? What do you do when you run into a roadblock and can’t get the code?
6. What kind of information do you usually capture about the video games/hardware you are emulating and why?
7. What do you consider an authoritative source for information about a video game? Clarifying questions: This includes information such as the game’s release date, the company that published it, the creators, and the country of origin if it’s a translation. If you can’t find any of this in the game itself, how do you find it?
8. If a game is known by multiple names (i.e., Final Fantasy VI in Japan was released as Final Fantasy III in North America), how do decide which version of the name to use? Do you treat each version as a different game?
9. Does copyright affect the kind of work you are able to do? If so, in what way?
10. How is the code of these games preserved?
11. Who do you think should be preserving games?
12. Do you think that anything else needs to be done to preserve these games?
13. Is there anything that I haven’t asked you that you’d like to talk about in relation to video game preservation?