Jasmyn R. Castro
New York University
Tisch, MIAP
Digital Preservation - Assignment #1
September 15, 2014

In a 2009, around the same time Archive Team began it's work, Dan Fletcher wrote a TIME article about the day Yahoo! pulled the plug on the popular web posting service known as GeoCities. "*The site was one of the first to offer home pages to the masses, letting users reserve a plot in a digital city...*" **[4]**. For one reason or another, Yahoo! decided to close the doors on GeoCities and discontinued registration to the site. It was at this time that Jason Scott, leader of Archive Team, took action and formed a small archiving crew that consisted of thirty volunteers. Their mission? To copy all the GeoCities webpages they could get their hands on before the site was shut down forever in April of 2009. Their stance was clear. "*When corporations take custody of your data, they not only have a responsibility to you, they also have one to humanity as a whole...*" and "*...in the long run, the question of what to keep and what to discard should not be made in corporate boardrooms*" **[3]**. Five years later, Archive Team seems to be even more committed to preventing the loss of web heritage.

Since 2009, Archive Team has been investing their precious time and energy into saving important bits of our digital heritage that most people don't end up missing until it's too late. When websites cease operations, merge with another website or company, shutdown or quite simply delete their data; Archive Team makes it their responsibility to save that data or content before it disappears forever. "*Archive Team is a loose collective of rogue archivists, programmers, writers and loudmouths dedicated to saving our digital heritage*" **[1]**. Their main focus is saving websites, which are born digital as far as design, but can often contain embedded media that has been digitized from other formats. For example, individual social media accounts often feature albums full of scanned family photos and video sharing sites like YouTube and Vimeo can feature digital video that may have been in a different format at one point in time. As far as their selection process is concerned, they simply ask themselves 'What out there, if it went away, would seriously wreck people?' **[2]**.

So, how does Archive Team actually save these websites? They call it *Deathwatch*. It is their process of keeping track of sites that appear to be gradually fizzling out. Archive Team also conducts what they call *Fire Drills*. This is a regular audit or check-in of a website that would not currently raise any red flags as far as expected sustainability is concerned, but holds a lot of the responsibility for saving the data and content of their site. On the Archive Team website, there is a list of all the sites they are currently attempting to download and save. Some of their current projects include TwitPic, Verizon customer pages, Quizilla, four genealogy centered websites that are merging with Ancestry [dot] com, and URLTeam.

Projects they recently concluded include Google Reader, Posterous, MobileMe, Tabblo, FanFiction [dot] net, Geocities, and FortuneCity. Once Archive Team has had a chance to go through the downloaded data and sort it out, they made most of the digital archives available for download on the *Archives* page of their website either as a direct download or a torrent file. They also upload these digital archives to the "Archive Team Collection" in the internet library known as the Internet Archive, where they are available for download as a zip file along with an optional

Jasmyn R. Castro
New York University
Tisch, MIAP
Digital Preservation - Assignment #1
September 15, 2014

XML file that contains pertinent metadata. While they have many success stories, not every undertaking is smooth sailing. For example, TwitPic, one of the first services that allowed Twitter users to attach photos to their posts, announced it was shutting down on September 25, 2014 **[4]** rather than battle Twitter in a trademark infringement lawsuit. When Archive Team started their TwitPic rescue, TwitPic founder Noah Everett blocked their access. On September 18, 2014, TwitPic announced that they were acquired by another company and would remain operational as a result.

Archive Team conducts what they call *Panic Downloads*; which are "*...full pulldowns of currently extant websites, meant to serve as emergency backups for needed sites that are in danger of closing, or which will be missed dearly if suddenly lost due to hard drive crashes or server failures*" **[1]**. This is extremely important because, under current law, no website is required to save users' data. On the Archive Team website, they have a section titled "Some Starting Points", where they provide some basic resources and archiving practices. Why? Because, aside from "*...serving as a hub for team-based pulling down and mirroring of data...*" they strive to provide "*...advice on managing your own data and rescuing it from the brink of destruction*" **[1]**. Right now, they offer the following **[1]**:

- *The Introduction* - an overview of basic archiving practices
- *Why Back Up?* - to put it bluntly, because they they don't care about you
- *Back Up your Facebook Data* - a guide to help individuals liberate their personal data from the grips of Facebook
- *Software* to assist with regaining control of individuals' data by providing tools for information backup, archiving and distribution
- *Formats* - a list of various formats to help familiarize people with the various data formats, and how to make sure files can be read later
- *Storage Media* - explains what it is, where to get it and how to use it efficiently
- *Recommended Reading* - a list of links to other websites for further information and research
- *Frequently Asked Questions* - Archive Team answers common questions

Jasmyn R. Castro
New York University
Tisch, MIAP
Digital Preservation - Assignment #1
September 15, 2014

Bibliography

[1] "Archive Team." *Welcome to Archive Team*. Archive.org, n.d. Web. 20 Sept. 2014.

[2] Bosker, Bianca. "Jason Scott's Archive Team Is Saving The Web From Itself (And Rescuing Your Stuff)." *The Huffington Post*, 27 Mar. 2013. Web. 20 Sept. 2014.

[3] Fletcher, Dan. "Internet Atrocity! GeoCities' Demise Erases Web History." *Time*. Time Inc., 09 Nov. 2009. Web. 22 Sept. 2014.

[4] Gohring, Nancy. "Lost Forever? Archive Team Says TwitPic Is Blocking Photo Downloads before Shutdown." *GeekWire*. N.p., 17 Sept. 2014. Web. 20 Sept. 2014.

[5] "Main Page." *Archiveteam*. N.p., 29 Aug. 2013. Web. 20 Sept. 2014.

[6] Sullivan, Mark. "The 'Archive Team' Rescues User Content From Doomed Sites." *PCWorld*. N.p., 12 Apr. 2012. Web. 20 Sept. 2014.