Appendix:

Archived Websites

- Bridges from Bamako. http://bridgesfrombamako.com/
  This is a blog site, created by Bruce Whitehouse, an anthropologist and former Peace Corps Volunteer. The blog is updated on a monthly basis (on and off). Nevertheless, this is one of the blogs we can rely on in terms of consistency. The blog includes an analysis of the 2012 political crisis in Bamako and it describes changes that have been taking place in Bamako since 2011. In addition, since the Mali crisis began in 2012, Whitehouse reports on some political issues as well. We ran our first test crawl for this blog site in mid-March. Our experience with the test was successful. However, running a real crawl was a different experience. Our first real crawl missed style sheets and images. For this reason, we had to troubleshoot, by scheduling patch crawls and setting up our crawls to ignore robots.txt file. After running a QA and patch crawls we were able to capture all content and make the blog live.

- MaliWatch. http://www.maliwatch.org/
  This is a non-political Non-Governmental Organization (NGO). The website lists non active projects but.
  We decided to capture content of this website quarterly. Our test crawl was successful, and we are happy to report that the real crawl was as well. We believe that this success is due to the content of the site. This site's content is heavy in pdf files and text files. Therefore, we did not have to capture many images and video content.

Central African Republic

- http://www.radiondekeluka.org/
  We experienced technical difficulties in capturing content of this site. Specifically CSS. The first real crawl did not grab all website content, specifically images. Therefore, we ran a patch and a QA report. After these problems were resolved, we had successful captures. This is one of the websites that we capture on a daily basis.

- http://www.centrafrique-presse.info/site/
  Centrafrique presse's primary goal is to inform the public on real time issues regarding policy, economy, human rights and social issues related to Central Africa.
  This is an example of a site for which we experienced ongoing capture issues on and off. It has become somewhat difficult to predict the outcome of captures, without close monitoring. Besides, the site does not have regular updates. For this reason, we decided to capture only three times since May, 2014. We still have some images that cannot be captured. Archive-It is working on finding a solution to this problem.

- http://www.centrafriquelibre.info/

Centrafrique Libre is a Central African independent newspaper created by many correspondents and world journalists with the aim of contributing to the freedom of press in Central Africa. We decided to crawl this website on a weekly basis. Test crawls ran successfully however; real crawls were not as successful. After troubleshooting a few times, we successfully gathered all content and we are currently archiving this newspaper site.

- http://centrafrique-presse.over-blog.com/
  Centrafrique Presse's blog was created with the main focus of informing the public on humanitarian aids and economic development issues in Central African Republic (CAR) and around the world. In addition, Centrafique-press promotes democratic governance and economic good for the people of CAR. This blog is updated on a daily basis (on and off) thus, we decided to crawl it daily. Test crawl was successful however, real crawl proved to be difficult as there was hard to capture content. After running a few patch crawls and ignoring robots.txt file, we are now able to capture all content.

- http://reseaudesjournalistesrca.wordpress.com/
  This is a blog administered by the Network for Journalists for Human Rights in Central African Republic (CAR-RJDH). RJDH is an organization consisting of journalists with a goal of defending and supporting human rights and humanitarian organizations in Central African Republic. This blog is updated on a weekly basis. Test crawl along with real crawl were successful. We continue to capture and store content for the blog.

DRC (Democratic Republic of Congo):

- Media Congo:
  http://www.mediacongo.net
  Media Congo is a news website which provides news from Congo and its diaspora. We decided to capture content for this site on a weekly basis. Test crawl was successful though, real crawls were not so successful. Since this website is rich in images and media, it proved to be difficult to capture all content without extensive troubleshoot and communication with Archive-It team. Design and layout of the website were part of the problem. Eventually, we captured all necessary content and we continue to do so.

- Infobascongo:
  http://www.infobascongo.net
  InfobasCongo is a non-profit organization comprised of computer and media professionals aiming to convey social-economic and political information about the DRC. This website is updated monthly, hence we crawl its content once a month. Our test crawls were successful however, the real crawls were not. Most images and video content required troubleshooting and working with the Archive-It team for a few weeks until we were able to successfully gather all content.

- Kleber - L'Observateur:
  http://kleber-lobservateur.blogspot.com

Kleber-L'observateur is a blog that highlights political, economic and social issues related to the African region while paying special attention to the Democratic Republic of Congo. The blogger focuses on regional political issues related to foreign policy.

We experienced similar problems as with the other sites. After a real crawl, we realized that some content was missing therefore, we needed to run a few patch crawls. After a patch crawl and QA, the website was successfully crawled. This blog is crawled on a monthly basis.

- Congo Siasa:
  http://congosiasa.blogspot.com
  This blog looks to bring to light the trials and tribulations of Congo and its people. The blogger focuses on the conflicts that occur in this area of the world as well as the tensions that exist between the government and the militia. These conflicts are also looked at in light of the interactions between the Congo and other countries around the world.
  Real crawl was unsuccessful. After troubleshooting a few times, we were able to capture all content. This blog is crawled on a monthly basis.

- Societe Civile.cd:
  http://www.societecivile.cd
  La Société Civile is a portal open to the world. It provides peace and democracy, human rights and elections news of the Democratic Republic of Congo (DRC). In addition, the portal offers regional resources from the DRC provinces and information related to health, education, business and trade, agriculture and livestock, nutrition, and social assistance. This website is updated monthly. Due to the richness of media content we experienced technical problems with captures. After running a few patch crawls and QAs we succeed.

- Churches, Peacebuilding and Women's Rights in DRC:
  http://www.drcongo.nibrinternational.no
  This blog is part of the ongoing research project Religious Civil Society Networks in the Great Lakes region. The project, funded by the Norwegian Ministry of Foreign Affairs is set to help in peace-building process, Peace and Security and in the implementation of the United Nations Security council Resolution 1325 on Women. The research institutions involved in this project are the Norwegian Institute for Urban and Regional Research, Centre for Intercultural Communication in Stavanger, the Evangelical University in Africa in Bukavu and Universite Officielle de Bukavu. Updates for the blog take place yearly. We have only crawled the site once. Crawling experience mirrors that of the other sites in terms of running a few patch crawls.

Newly added site:

- Friend of Liberia
  http://fol.org/
  Friends of Liberia (FoL) is a non-governmental, non-profit organization which supports Liberia through socio-economic, humanitarian and education advocacy programs.
  Crawling this site was similar to the other sites above. The site has a marquee with videos, hence, we had to troubleshoot. After troubleshooting, images and video were captured and preserved. FoL is captured monthly.

Problem sites:

- Radio Okapi:
  http://radiookapi.net
  Radio Okapi is part of the United Nations in the DRC. We were able to capture some content of this website. This radio station was established with the main goal of spreading independent national, regional and international news in a rigorous, factual and credible form. Content of the website is updated weekly. After we had captured content for three months unexpectedly, we started experiencing problems with the display of images and video captures. After troubleshooting the problem could still not be resolved. For this reason, we have asked the web administrator to unblock some content.

- Le Republicain
  http://lerepublicain-mali.com/
  This website covers Mali's recent topics including social, cultural, technological and environmental issues. In, addition it links its audience to external websites such as the UNICEF. Unfortunately, this newspaper website has been disabled and online access was terminated. Therefore, we stopped capturing its non-existing content. Fortunately, we have two months' worth of content preserved.

- Mouvement National de Libération de l'Azawad, MPLA. http://www.mnlamov.net/
  The National Movement for the Liberation of Azawad (MNLA) blog, was created with the intention of alerting the international organizations for human rights protection to the abuses of human rights by the Malian army on civilians.
  We set up this site to be crawled quarterly. However, we experienced technical issues capturing videos from the beginning. The main page of the site is composed of a video marquee which proved to be difficult to capture. There seem to be strange redirects that remove parts of the URL. Archive-It has been informed of the problem since Sept 08. They last reported that engineers are working on finding a fix. For this reason, we decided to omit this website from public view, until the problem is resolved.