Collection Building

Collecting and preserving the Ukraine conflict (2014-2015): a web archive at University of California, Berkeley
Liladhar R. Pendse

Article information:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Collecting and preserving the Ukraine conflict (2014-2015): a web archive at University of California, Berkeley

Liladhar R. Pendse
Doe Library, University of California Berkeley, Berkeley, California, USA

Abstract

Purpose – The purpose of this paper is to highlight the web-archiving as a tool for possible collection development in a research level academic library. The paper highlights the web-archiving project that dealt with the contemporary Ukraine conflict. Currently, as the conflict in Ukraine drags on, the need for collecting and preserving the information from various web-based resources with different ideological orientations acquires a special importance. The demise of the Soviet Union in 1991 and the emergence of independent republics were heralded by some as a peaceful transition to the "free-market" style economies. This transition was nevertheless nuanced and not seamless. Besides the incomplete market liberalization, rent-seeking behaviors of different sort, it was also accompanied by the almost ubiquitous use of and access to the internet and the internet communication technologies. Now 24 years later, the ongoing conflict in Ukraine also appears to be unfolding on the World Wide Web. With the Russian annexation of Crimea and its unification to the Russian Federation, the governmental and non-governmental websites of the Ukrainian Crimea suddenly came to represent a sort of "an endangered archive".
Design/methodology/approach – The main purpose of this project was to make the information that is contained in Ukrainian and Russia websites available to the wider body of scholars and students over the longer period of time in a web archive. The author does not take any ideological stance on the legal status of Crimea or on the ongoing conflict in Ukraine. There are currently several projects that are devoted to the preservation of these websites. This article also focuses on providing a survey of the landscape of these projects and highlights the ongoing web-archiving project that is entitled, "the Ukraine Crisis: 2014-2015" at the UC Berkeley Library.
Findings – The UC Berkeley's Ukraine Conflict Archive was made available to public in March of 2015 after enough materials were archived. The initial purpose of the archive was to selectively harvest, and archive those websites that are bound to either disappear or change significantly during the evolution of Crimea's accession to Russia. However, in the aftermath of the Crimean conflict, the ensuing of military conflict in Ukraine had forced to reevaluate the web-archiving strategy. The project was never envisioned to be a competing project to the Ukraine Conflict project. Instead, it was supposed to capture complimentary data that could have been missed by other similar projects. This web archive has been made public to provide a glimpse of what was happening and what is happening in Ukraine.
Research limitations/implications – Now 24 years later, the ongoing conflict in Ukraine also appears to be unfolding on the World Wide Web. With the Russian annexation of Crimea and its unification to the Russian Federation, the governmental and non-governmental websites of the Ukrainian Crimea suddenly came to represent a sort of "an endangered archive". The impetus for archiving the selected Ukrainian websites came as a result of the changing geopolitical realities of Crimea. The daily changes to the websites and also loss of information that is contained within them is one of the many problems faced by the users of these websites. In some cases, the likelihood of these websites is relatively high. This in turn was followed by the author's desire to preserve the information about the daily lives in Ukraine's east in light of the unfolding violent armed conflict.
Originality/value – Upon close survey of the Library and Information Sciences currently published articles on Ukraine Conflict, no articles that are currently dedicated to archiving the Crimean and Ukrainian situations were found.

Keywords Conflict, Collection building, Crimea, Endangered archive, The Ukraine crisis, Web-archiving

Paper type Case study

Introduction and problem statement

The Euromaidan civil protests that began in Kiev in November 2013 aimed at the closer integration with the European Union and were a sort of culmination of expression of popular sentiments of those who participated in them against the cronyism, corruption and lack of economic growth in Ukraine over the period of time. The election of the then President Yanukovich and his subsequent exit from the power were also fueled by the extensive use of the social media by the Ukrainian activists. Besides the opposing viewpoints, there were multiple websites that were providing, often times, conflicting information to their audiences. These websites varied from individual blogs, online news media, Facebook pages and also local governmental sites. The further continuation of the conflict meant multiple viewpoints, multitude of the differing opinions and the discerning response to the conflict by both governmental and non-governmental authorities on the web.

There were several reasons behind our project, one of them being capturing and preserving multiple narratives that emerged after the post-Euromaidan conflict in Ukraine (Bachmann and Lyubashenko, 2014). These narratives have continued to evolve to the day of writing this article. The purpose of this project was also to create, by preserving the websites that might eventually be lost to the ravages of time, political position taking (Ivshyna and Den', 2014). Besides the Euromaidan protests, there were multiple watershed moments in the current history of Ukraine. The Ukrainian conflict acquired different significance with the annexation of Crimea in March 2014 by Russia despite the agreement about maintaining the territorial integrity of Ukraine (Budapest Memorandums on Security Assurances, 1994). The ensuing conflict in the Eastern part of Ukraine and the shooting of the Malaysian Airlines flight MH-17 in July 2014 not only heightened the attention to the "chatter" about shooting down of the "military transport" by the rebelling forces, but also internationalized the Ukrainian conflict further. The repercussions of these events can also be understood as the narratives of the disagreement. The disagreement among the Ukrainians in various spheres of social engagement meant these narratives finding their reflection not only on the battlefields, press or television, but also in the different facets of the World Wide Web.

These disagreements were expressed not only on the social media platforms like VKontakte, Moi Mir, Facebook and blogs, but also in various online news reporting sites. In case of the governmental websites like those like in Crimea, the outright physical takeover of the administration meant several governmental websites that once were managed by the Ukrainian authorities were bound to be dysfunctional or whose content management became questionable. The further evolution of the military conflict, the establishment of so called "Donetsk People's Republic (DNR)" and "Lugansk People's Republic (LNR)" by the separatists showed a need for the strategy to preserve these websites. This uncertainty in the future continuation of a particular website or its historical content served as an impetus for the website preservation project at the University of California. In this article, I will discuss the selection criteria for the websites that are currently being archived. I will expound on the scope of this archive and describe its contents. Lastly, I will compare the UC Berkeley's Ukraine conflict archive with other web-based initiatives that currently are performing the similar archiving of the Ukraine conflict.

Literature review

The web archiving of the various websites has been discussed extensively in current Library and Information Science-related literature (Masanès, 2006a; Kavcic-Colic and Klasinc, 2011). The earliest concerns about the need for web archiving in light of the sheer size of the new web-based resources and their short web life were geared toward preserving what is possible. Various motivational factors behind the web as a platform for providing information have been discussed by Berners-Lee and Fischetti (1999). In 2008, Pinsent in the Preservation of the Web Resources Handbook has analyzed the motivations behind the web archiving. These motivations vary from need for providing access to the activities such as annual reports,

limited financial data as an evidence for a particular organizations existence to preserving the scientific, cultural and intellectual output of a particular organization (Pinsent *et al.*, 2008). This means that web archiving as a preservation technique can only preserve a limited portion of the web. The choices thus have to be made consciously about which websites one would choose to preserve as a part of the archive.

On one hand, the web archiving has been framed by some as limited, directed preservation of those websites that are endangered, while on the other hand, the web archiving has been understood as a part of the "open-publishing" system (Masanès, 2006b). As far as the selection of the websites for archiving is concerned, the thematic approach to the subjects or grouping these websites by the format can be of utility. This arrangement of an archive is usually independent of the purposes behind such archiving with one exception. The exception is to make the content available when it does not exist in real time to the user.

The purposes of web archiving have varied from archiving personal websites to preserving the contents of the institutional repositories or governmental agencies (Phillips and Koerbin, 2006; Sinn and Syn, 2014). However, one should differentiate among the archiving of the personal information as it is submitted to the social networking sites and directed and purposive harvesting of data along with its long-term preservation for specific use of a particular subject. There have been numerous studies on the at-risk websites and their life cycles.(Dell and Shultz, 2014; Bearman, 2006). Irrespective of the contexts of the subject matters of these studies, these studies have alluded to the fact that for the recordkeeping purposes, the long-term preservation of the websites becomes of importance. There have been also ongoing archiving efforts of the web of governmental information (Gomes *et al.*, 2011). The current state of web archiving has been summarized by Meghan and Mayer in their 2014 article entitled "Community, tools, and practices in web archiving: The state-of-the-art in relation to social science and humanities research needs" (Dougherty and Meyer, 2014).

The questions about the criteria for preserving certain websites over the others have been discussed in the recent Archival Studies and Library and Information Studies literature. In 2003, the UNESCO released its guidelines on preserving the digital heritage for posterity (National Library of Australia, 2003). The Library of Congress' collections policy statements supplementary guidelines states the following:

> Since its inception, web archiving at the Library has primarily been a collection-based activity. This means that the usual practice is not to acquire individual web sites one-by-one, but as part of a named subject, event, or theme-based collection. The sites harvested for the collection are curated by Recommending Officers (ROs), who set the frequency and scope of the harvesting of a site. The Library's goal is to create an archival copy – essentially a snapshot – of the site at a particular point in time or over a period of time. (Library of Congress, 2016)

The philosophical discussion about what constitutes and what does not constitute the digital heritage when it comes to preserving the Ukrainian conflict-related websites is out of the scope of this paper. The websites with differing viewpoints are over time bound to:
- change their contents; or
- become extinct.

It is sufficient to state that the ephemeral nature of information that is displayed in these specific websites means that there is an increased need for the preservation of these websites. The needs for digitally preserving and archiving these websites have been recognized under the auspices of the several projects.

These projects are well-described on a web page of the Association of Slavic, East European and Eurasian Studies (Zitser, 2014). These projects have varied from creation of a LibGuide at Duke, or a web page with the different Ukraine-related links by the UCL School of Slavonic and East European Studies Library to a web-archiving project called Archive-It-Ukraine Conflict. The first two guides seem to be rather static at the time of the writing of this article. The page of the UCL was last updated on September 10, 2014 at the time of writing this article (Ukrainian Crisis, 2014). Archive-It-Ukraine Conflict was established by the Internet Archive in collaboration with the specialists from Library of Congress, Stanford's Hoover Institution, the Harvard Ukrainian Research Institute and the Global Investigative News Network. Besides the Archive-It project and "Ukraine Conflict Archive" at UC Berkeley, I could not find any evidence of any other web-archiving project that was publicly available.

Web-archive platform and choices of the websites

The University of California's Web Archiving Service (WAS) has been described by the author in his earlier article (Pendse, 2014). It was established in 2007 and the WAS functions as the information storage and retrieval system for the websites that have been "crawled and preserved". The WAS allows its user to create and manage simultaneously multiple projects. It allows users to crawl and cull data from a particular website for brief 1-h interval or more detailed 36-h capture. The WAS allows its user to capture also the linked data if needed. The WAS also allows the user to compare the data from multiple captures. During the comparison of two or more captures, the user can tell how many new files have been added. How many files that were previously in the earlier capture have been replaced with the new files? What are their formats and what are their sizes. The other important features are automated captures by daily, weekly, monthly and custom capture options and ability to browse the archive using the WAS browser.

The WAS service gives the project manager capability to override the robots.txt file, thus ensuring the capture of the information that is contained in a publicly available website. In January 2015, the California Digital Library (CDL) announced that the WAS will partner with the Archive-It in the summer of 2015 for crawling, managing and creating new tools for the purpose of the web archiving (California Digital Library (CDL), 2015). The project "The Ukraine Crisis: 2014-2015" was conceived in early March 2014, independently of the current Archive-It project that was mentioned earlier. The criteria for the preservation were aligned with those that the Library of Congress recommends under the acquisitions collections supplementary policy. These criteria were based on usefulness of the information that these websites might contain in serving the current and future information needs of the user of such archive. Second, the information that is contained within these sites in unique and is at risk. Third, such a collection of information and preservation of these websites falls within the domain of the curatorial collection development responsibilities.

The impetus for this project at the University of California Berkeley lay in the fact that the Ukraine crisis had become a part of our hot topics-related library guide (Pendse, 2014a). During the creation of the guide, I realized that there wasn't any way to preserve the constantly changing information on various Ukraine-related websites as the conflict progressed. The LibGuide provided a static platform that was conducive for conveying the basic information that could be used for instructional purposes; however, its utility was limited when it came to presenting the evolving web-based information.

Based on the previous experience of archiving the endangered lesbian, gay, bisexual and transgender (LGBT) websites of Eastern Europe, initially I had proposed in detail a pilot project about archiving the websites of Ukrainian conflict similar to the one I had initially undertaken in 2013 (Pendse, 2014b). However, in consultation with the Associate University Librarian and the Director of Digital Initiatives and Collaborative Services Dr Erik Mitchell, I decided to launch a more substantial project. It was envisioned that under this project the Ukrainian conflict archive will take in consideration the needs of not only the local scholars, but also of those who normally will not have access to the archived information. At the point of the decision making in March 2014, the author was not aware of the parallel efforts that were taking place at Stanford University Library in conjunction with a multi-institutional collaboration through Archive-It.

The choices of the websites for the archiving project were discussed with several colleagues on campus that included the faculty members, the librarians who managed the digital projects in the library as well as the Area and International Studies colleagues. Initially, there were 20 sites that were chosen for archiving with the stipulation that I need to add more sites as the project progressed depending upon the changing geo-political circumstances in Ukraine. In 2015, as the conflict continued, the author identified and added nine additional sites. Before the capture of each site, a letter indicating the consent to capture the contents was sent to the address where it was available to seek the permission. To my regret, to the date of writing the article, I have failed to get a response from the managers of these websites. However, the web archive that has been created is based on the fair academic use clause and it is available to the public.

These sites were as indicated in Table I. To choose the sites, I used the purposive sampling strategy. At the initial stages, I focused more on archiving the online news-related websites. The emphasis was placed on archiving materials mostly in non-English languages with an exception that some English language websites that belonged to the Ukrainian entities will be archived. For the personal websites of the individuals or their supports who emerged in 2014 as some of the key players in this conflict, I used an article that was published by the BBC in August 2014 (BBC News, 2014). The sites were chosen to represent differing viewpoints on the actors in this conflict.

Only in limited cases, the captures were automated on frequencies varying from weekly to monthly. However, the

Table I The websites that are the part of the web-archive on Ukraine conflict: 2014-2015

Title of the website	Hyperlink	Genre	Language of capture	No. of captures	Last capture
Agentstvo Novorossiia	www.novorosinform.org/	News Agency/Information Portal	Russian	3	2/20/2015
Aktual'naia Informatsiia	http://etoonda.livejournal.com/	Blog	Russian	1	2/11/2015
Antimaydan-LNR (1)	http://antimaydan.info/luganskaya_respublika_lnr.html	News-Lugansk	Russian	3	2/20/2015
Cyber-Berkut	http://cyber-berkut.org/en/	Pro-Russian Hacktivist	English	3	2/15/2015
Euromaidan Press	http://euromaidanpress.com/	Online Newspaper	English	2	2/15/2015
Informatsionnoe Soprotivlenie	http://5protyv.info/ru/o-nas	Ukranian News Analysis	Russian	1	2/15/2015
Luganskaia Narodnaia Respublika: Respublikanskii Informatsionnyi Portal (1)	http://lnr-portal.su/	News Agency/Information Portal	Russian	2	2/15/2015
Mariupol'skie Novosti	www.mariupolnews.com.ua/	Online Newspaper	Russian	4	2/15/2015
Novosti Novorossii	http://novorospolk.ru/	News Portal	Russian	2	2/15/2015
Official Blog of Norossinform	http://novorosinform.livejournal.com/	Social Media	Russian	1	2/15/2015
South Front	www.youtube.com/user/ygfront and www.southfront.eu/	Social Media	English	2	2/15/2015
Strelkov on Vkontakte	https://vk.com/strelkov_info	Social Media	Russian	2	2/15/2015
112.ua	http://ua.112.ua/http://112.ua/	News Agency/Information Portal	Ukranian/Russian	25	2/15/2015
Den' Kiev (English Edition)	www.day.kiev.ua/en	Online Newspaper	English	27	2/15/2015
Donetskie Novosti	http://dnews.donetsk.ua/	Online Newspaper	Russian	33	2/15/2015
The Government Portal-Ukraine (English)	www.kmu.gov.ua/control/en	Government Portal	English	2	2/15/2015
The Government Portal-Ukraine	www.kmu.gov.ua/control/	Government Portal	Ukranian	107	2/15/2015
Grazhdanskaia Oborona: Khronika Euromaidana	http://911sevastopol.org/	Online Newspaper	Russian	1	4/9/2014
Kiev Agency for Investigative Reporting	http://slidstvo.info/	Online Newspaper	Ukrainian	12	7/22/2014
Krymska Svitlitsa	http://svitlytsia.crimea.ua/	Online Newspaper	Ukrainian	24	5/28/2014
Kyiv Post	www.kyivpost.com/	Online Newspaper		3	2/15/2015
KUKIiT	www.uncat.crimea.ua/ua/?l=ua	University Website	Russian	23	5/29/2014
Nikolaevskii Tstentr	http://nikcenter.org/	News Portal	Russian	105	2/15/2015
Perekop: Everything about Crimea.	www.perekop.info/	Crimean Information Portal	Russian	11	5/31/2014
Slavgorod.com.ua	http://slavgorod.com.ua/	News	Russian	12	2/15/2015
Gazeta Vechernii Donetsk	http://vecherka.donetsk.ua/	Online Newspaper	Russian	6	2/15/2015
Gosudarstvennyi Sovet Respubliki Krym	Past: www.rada.crimea.ua/ Current: www.crimea.gov.ru/	Government Portal-Crimea	Russian	13	6/1/2014
Munitsipal'naia Gazeta-Donetsk	http://mungaz.net/	Online Newspaper	Russian	6	2/15/2015
Crimea-Fiscal Office	http://crimea.sfs.gov.ua/	Government Portal	Ukrainian/Russian	24	5/28/2014

majority of the websites that constitute a part of this archive were on a one-time manual archiving mode. This blended method for archiving was chosen simply because I wanted to monitor the space that was being used judiciously. The goal was to capture important vignettes of certain events as the political and military activities unfolded in Ukraine.

As of March 1, 2015, there were 478 captures completed. Figure 1 below shows the landing page of the project.

The WAS allows the project manager to deactivate certain sites and reactivate them as needed. This way, the project manager can focus on manually archiving only the limited numbers of active sites. As of February 2015, there were only 13 sites active (see Figure 2).

The manage site feature of the project allows the project manager to search for specific captures. The platform allows also for comparing several captures of the same website by identifying the unique sites.

As discussed, the strategy of limited and incidental archiving allows the project manager to generate savings of the storage place. The gathered data and the program's archiving patterns can be analyzed using the watershed moments in the conflict. First, it was the annexation of Crimea by Russia. Second important event was the shooting down of the Malaysian Airlines MH 17 flight on July 17, 2014. Third important event was the beginning of the "Normandy" rounds in Minsk. And the fourth important event was the

Figure 1 The interface of the WAS

Figure 2 Manage sites
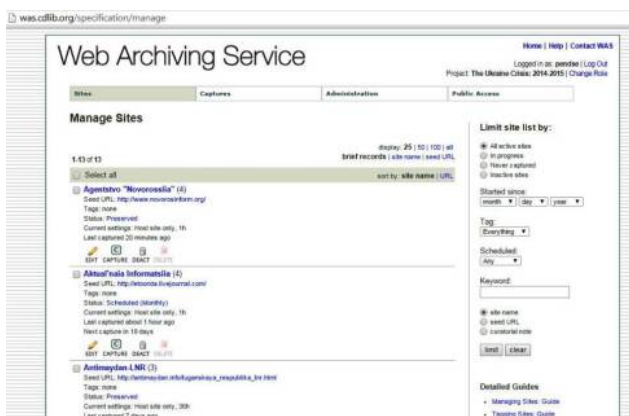
assassination of Boris Nemtsov in Moscow on February 27, 2015. While the assassination of the opposition member might not directly impact the outcome of the Ukraine conflict, it will be interesting to see to what extent was discussed in Ukrainian press both in the Western and Eastern parts of country. Boris Nemtsov was slated to take part in rally.

The Ukrainian conflict remains much more complicated than rather simplistic depiction of its progress in stages. The using the stages to capture and preserve the data from the selected websites provides one way to organize the "archive" while being cognizant of the resource constraints that are needed to preserve these captures. Also, numbers of the captures of a particular site were determined by the dates when such sites were identified and added to the archive. In some cases, the site capture was automated for monthly captures based on the relative importance of the site in the archive.

A brief description of the select sites in the archive.

Agentstvo novorossiia (www.novorosinform.org/)

By self-description, this site is the "official" news agency of the "Federal Republic of Novorosiia" (Агентство "Новороссия" – О редакции, 2015). This site reports the news from the rebels in the Ukraine's east. The site's domain registration is protected in the WHOis.net.

Aktual'naia informatsiia (http://etoonda.livejournal.com/) is a Russian blog site that sympathizes with Ukrainian rebels and provides information on the events of the political significance in Ukraine's East. The site's name is made of up three Russian words eto on da. This can be translated as It (is) him yes. The site also embeds various relevant YouTube clips. However, the last capture shows that the blog has not been updated since April of 2014 despite the fact that profile states that the blog was updated in April of 2015.

Antimaydan-LNR

This site is apparently managed either by the "Lugansk Peoples Republic" or its sympathizers. Lugansk or Luhansk is a city in Eastern Ukraine that was known as Voroshilovgrad during the Soviet times. In May 2014, a referendum was held to assert Luhansk's "independence" from the Central Ukrainian authorities and on 25 June of the same year, the city of Lugansk was proclaimed as the capital of this "breakaway" region. However, to this day, the Lugansk People's Republic remains unrecognized by any of the world's countries (Figure 3).

The whois database search reveals that the domain is registered to a private individual, and thus, one cannot comment on the official nature of the website of the self-proclaimed Lugansk republic. However, the information on the site is updated regularly, and the contact feature of the website as depicted below listed several phone numbers at the time of writing of this article (Figure 4).

The archive also includes websites in English language. During the envisioning of the archive, a conscience decision was made to include sites in English language along with Ukrainian and Russian languages. One such time that forms an integral part of this archive is Euromaidan Press (http://euromaidanpress.com/).

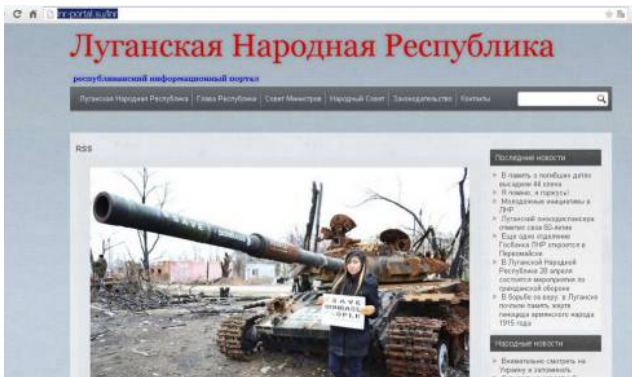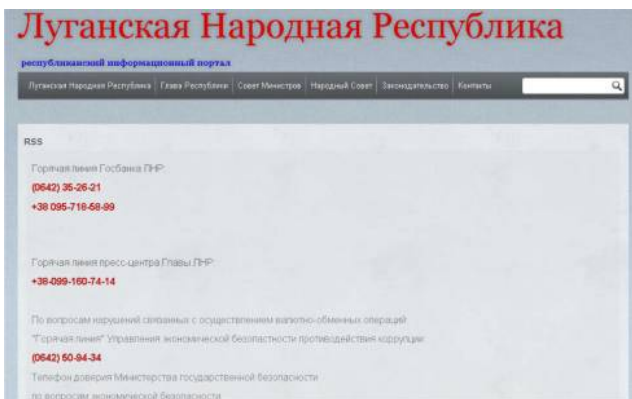**Figure 3** Antimaydan-Lugansk people republic

Figure 4 The contact information for the Antimaydan-LNR website

Euromaidan press

The site's self-description is as follows:

> Euromaidan Press is an online newspaper founded in 2014. As a media outlet, EP focuses on news and events in and relating to Ukraine, and is driven by articles both from professionals and experts, as well as volunteer contributions. Specializing in translations of local Ukrainian news outlets, EP strives to be the go-to bridge between Ukraine and the English-speaking world.

Besides English, the site has German language option (Figure 5).

As one can note from the choices of sites that are included in archive consists primarily of several media sites, nevertheless there are several governmental sites including some crawls of the Governmental Portal of Ukraine.

The WAS interface and the analysis of a crawl

In this section, I wanted to highlight some specific features of the WAS interface that demonstrate ease of its use as well as its organization. The interface of the WAS is intuitive and it allows its user to work on multiple projects at the same time as depicted below. Each project is maintained separate from the other for the ease of the curator (Figure 6).

Here, I am shown as an administrator/curator for three different projects. The interface of the project-specific page allows a curator to create a new website, manage the existing website that has been archived, view the captures and more importantly to compare various captures of the same resource.

Figure 5 The landing page of Euromaidan press

Figure 6 The WAS projects landing page

The captured websites are arranged alphabetically (Figure 7).

The interface allows a curator to further see each capture and to compare it to the previous captures. For example, one sees below in the picture the details of captures of Antimaydan-LNR (Figure 8 and 9).

There are four separate captures for the website called "Antimaydan-LNR". The change comparison interface is depicted in the picture below (Figure 10).

For the purpose of this paper, to demonstrate the WAS capabilities, I have chosen two captures of the website (Figure 11).
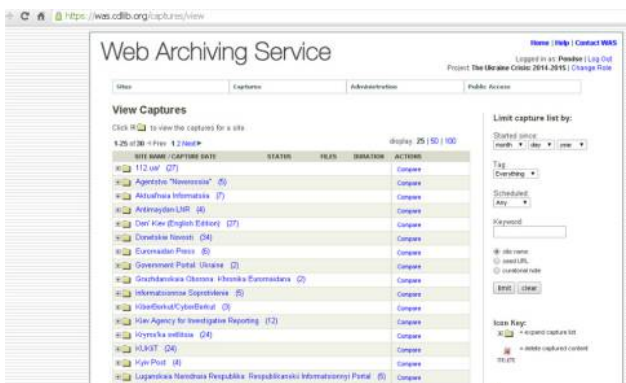
Figure 7 The list of captured websites

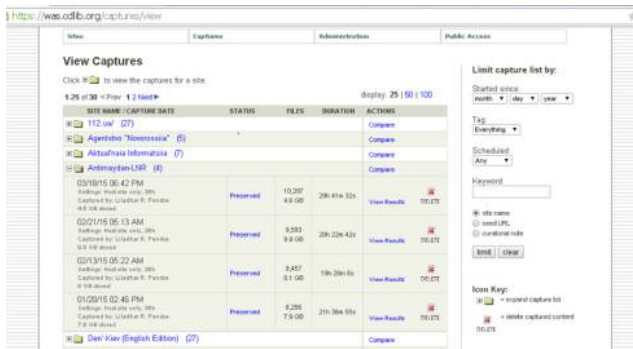**Figure 8** The details of the captures of Antimaydan-LNR

Figure 9 A capture of the Antimaydan-LNR
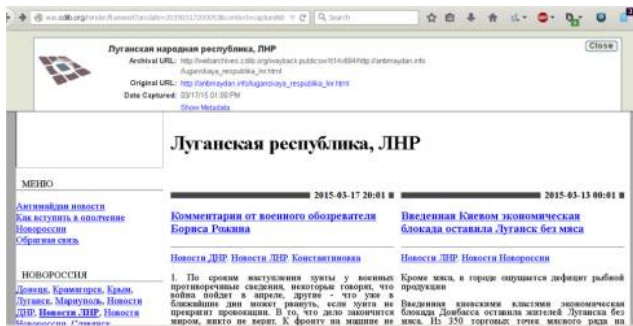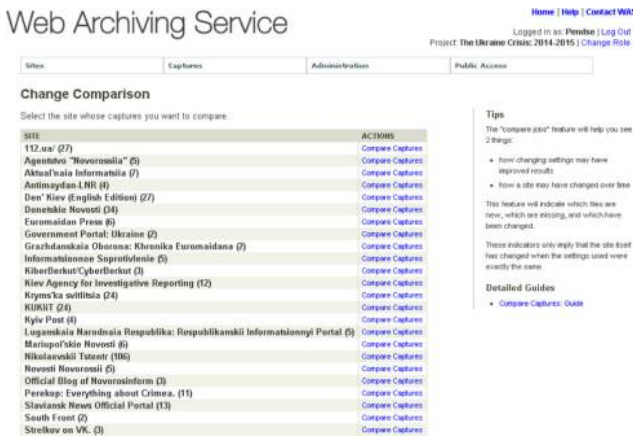

Figure 10 Change comparison

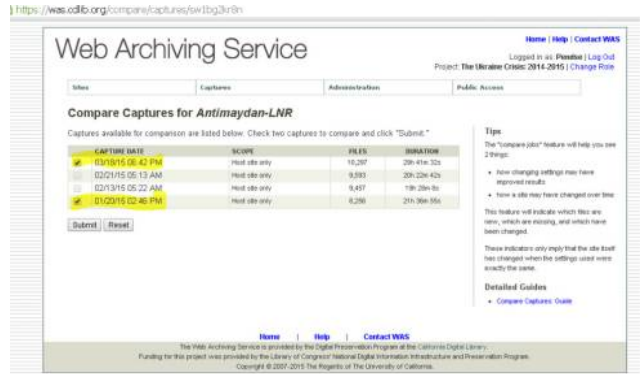**Figure 11** The capture comparison interface

Figure 12 Comparison processing progress

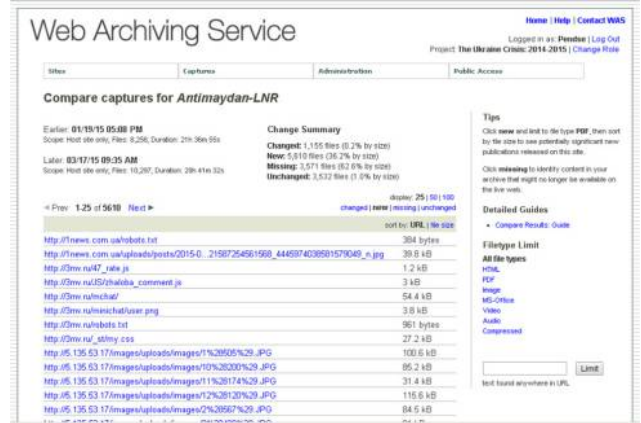**Figure 13** The changes in captures by the types of files

Figure 14 The WAS reports interface

This comparison allows us to see which files of a particular website have changed and evolved (Figure 12 and 13).

We can sort the files by the types such as HTML, PDF, Image, etc.

Besides the crawl comparisons, the WAS allows its user to run detailed reports (Figure 14).

While the web-crawling reports are of technical nature, these allow curator about the certain insights about the nature of the structure of the websites that were crawled. In case of need, technical experts can troubleshoot problems that might be associated with these crawls and web archiving (Figure 15).

Figure 15 The crawl report

Discussion and comparison with Archive-It-Ukraine Conflict.

The Ukraine Conflict project at the Archive-It began one month earlier that our project at UC Berkeley Library. Since February 2014, the project has sought to:

> [. . .] to document conflict in Ukraine as it progresses. Contributions to this collection were made by the Archive-It team and subject matter experts from Stanford University, Library of Congress, Global Investigative News Network, and Harvard Ukrainian Research Institute in the fields of Investigative Journalism, Russian, and Eurasian studies. Content includes news outlets, social media, blogs, and government websites. Sites are in English, Ukrainian, Russian, and other languages. (Archive-It, 2014)

Currently, this archive includes 71 websites, 99 articles, 19 blogs, 11 Social Media and 5 Facebook pages along with other formats. This archive is a collaborative effort of Library of Congress, Harvard Ukrainian Research Institute and Hoover Institute of Stanford's with the Internet Archive. In scope, this archive is larger when compared in its sheer size and the diversity of the websites and other media that it has captured (Figure 16).

The Archive-It project is hosted on the Wayback Machine platform. While the UC Berkeley's Ukraine Conflict project is based on the CDL's WAS platform. Both platforms enable the curators in archiving the websites and making them available for the posterity. Below, you can see the capture data of the Internet Archive's crawl (Figure 17).

The user can thus look up the crawl by date. In this particular case, we see a single crawl. When one clicks on the date's hyperlink, one gets the archived web page as depicted in Figure 18.

Figure 16 Archive-It: Ukraine conflict

Figure 17 Capture data of the Ukraine conflict Web archive

Figure 18 The archived website

Conclusion

The UC Berkeley's Ukraine Conflict Archive was made available to public in March 2015 after we had archived enough materials[1]. The initial purpose of our archive was to selectively harvest, and archive those websites that are bound to either disappear or change significantly during the evolution of Crimea's accession to Russia. However, in the aftermath of the Crimean conflict, the ensuing of military conflict in Ukraine had forced to reevaluate our web-archiving strategy. The project was never envisioned to be a competing project to the Ukraine Conflict project. Instead, it was supposed to capture complimentary data that could have been missed by other similar projects. This web archive has been made public to provide a glimpse of what was happening and what is happening in Ukraine.

Note

1 Available at: www.cdlib.org/cdlinfo/2015/04/24/was-service-update-march-2015/

References

Archive-It (2014), Available at: https://archive-it.org/collections/4399

Агентство «Новороссия» - О редакции (2015), Агентство «Новороссия» - О редакции, July 1, 2014, available at: www.novorosinform.org/about (accessed 28 February).

Bachmann, K. and Lyubashenko, I. (2014), *The Maidan Uprising, Separatism and Foreign Intervention: Ukraine's complex transition*, Peter Lang GmbH, Internationaler Verlag der Wissenschaften, Frankfurt.

BBC News (2014), "Key players in East Ukraine Unrest", *BBC News*, 24 August, available at: www.bbc.com/news/world-latin-america-27211501 (accessed 11 February 2015).

Bearman, D. (2006), "Moments of risk: identifying threats to electronic records", *Archivaria*, Vol. 62, pp. 15-46, available at: http://search.proquest.com/docview/57674568?accountid=14496

Berners-Lee, T. and Fischetti, M. (1999), *Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web by its Inventor*, Harper San Francisco, San Francisco, CA.

Budapest Memorandums on Security Assurances (1994), "Council on foreign relations", *Council on Foreign Relations*, 5 December, available at: www.cfr.org/nonproliferation-arms-control-and-disarmament/budapest-memorandums-security-assurances-1994/p32484 (accessed 3 February 2015).

California Digital Library (CDL) (2015), "Announcing a new partnership: California digital library, UC libraries, and internet archive's archive-it service", California Digital Library (CDL), 14 January, available at: www.cdlib.org/cdlinfo/2015/01/14/announcing-a-new-partnership-california-digital-library-uc-libraries-and-internet-archives-archive-it-service/ (accessed 7 February 2015).

Dell, E.Y. and Shultz, S.M. (2014), "Conserving digital resources: issues and future access", *Journal of Electronic Resources in Medical Libraries*, Vol. 11 No. 3, pp. 124-133, doi: http://dx.doi.org/10.1080/15424065.2014.937657.

Dougherty, M. and Meyer, E.T. (2015), "Community, tools, and practices in web archiving: the state-of-the-art in relation to social science and humanities research needs", *Journal of The Association For Information Science & Technology*, Vol. 65 No. 11, pp. 2195-2209, Academic Search Complete, Web 4 February.

Gomes, D., Miranda, J. and Costa, M. (2011), "A survey on web archiving initiatives", *Lecture Notes in Computer Science*, pp. 408-420.

Ivshyna, L. and Den' (Kiev, Ukraine) (2014), *Liudy Maĭdanu, Khronika*.

Kavcic-Colic, A. and Klasinc, J. (2011), "Web archiving in the national and university library", *Knjiznica*, Vol. 55 No. 1, pp. 209-232, available at: http://search.proquest.com/docview/1266143501?accountid=14496

Library of Congress (2015), "Collections policy statements and supplementary guidelines", Acquisitions (Library of Congress), Library of Congress, available at: www.loc.gov/acq/devpol/cpsstate.html (accessed 11 February 2015).

Masanès, J. (2006a), *Web Archiving*, Springer, Berlin.

Masanès, J. (2006b), "Web archiving: issues and methods".

National Library of Australia (2003), *Guidelines for the Preservation of Digital Heritage*, National Library of Australia, Canberra.

Pendse, L.R. (2014a), "Archiving the Russian and East European Lesbian, Gay, Bisexual, and Transgender 'a pilot project'", *Slavic and East European Information Resources*, Vol. 15 No. 3, pp. 182-196.

Pendse, L.R. (2014b), "Ukraine crisis: 2014-2015: article databases", UC Berkeley LibGuides, *LibGuides*, 1 March, available at: http://guides.lib.berkeley.edu/c.php?g=33959&p=216351 (accessed 7 February 2015).

Phillips, M. and Koerbin, P. (2006), "PANDORA, Australia's web archive: how much metadata is enough?", *Journal of Internet Cataloging*, Vol. 7 No. 2, pp. 19-33.

Pinsent, E., Davis, R., Ashley, K. and University of London (2008), *PoWR: The preservation of Web Resources Handbook: Digital Preservation for the UK HE/FE Web Management Community*, University of London Computer Centre, London.

Sinn, D. and Syn, S.Y. (2014), "Personal documentation on a social network site: Facebook, a collection of moments from your life?", *Archival Science*, Vol. 14 No. 2, pp. 95-124, doi: http://dx.doi.org/10.1007/s10502-013-9208-7.

Ukrainian Crisis (2014), *UCL Library Services: UCL SSEES Library: Ukrainian Crisis*, UCL Library, available at: www.ssees.ucl.ac.uk/library/directory/ukraine2014.htm (accessed 5 February 2015).

Zitser, E. (2014), "Online guides to resources about the crisis in Ukraine", *Online Guides to Resources about the Crisis in Ukraine*, ASEEES, 16 March, available at: http://aseees.org/news-events/aseees-blog-feed/online-guides-resources-about-crisis-ukraine (accessed 5 February 2015).

About the author

Liladhar R. Pendse is the collection development librarian for Slavic, East European and Eurasian Studies at UC Berkeley. Prior to coming to UC Berkeley's Doe Library in 2012, he has worked at Princeton University and UCLA libraries in the same capacities. He holds a PhD in Information Studies from UCLA. At Berkeley, he focuses on further evolving collections in multiple areas under his current curatorial responsibilities. These areas are Sub-Saharan Africa (on Interim basis), Armenian and Caucasus Studies, Central Asian and Mongolian Studies. Liladhar R. Pendse can be contacted at: Lpendse@library.berkeley.edu