

# WS-DL's Work towards Enabling Personal Use of Web Archives

**Michele C. Weigle**, @weiglemc

Web Sciences and Digital Libraries (WS-DL) Group, @WebSciDL

Department of Computer Science

Old Dominion University

December 18, 2018 / Library of Congress

---



# ODU WS-DL Group

## Faculty

- Dr. Michael L. Nelson
- Dr. Michele C. Weigle
- Dr. Sampath Jayarathna
- Dr. Jian Wu

## Graduate Students

- Scott Ainsworth
- Sawood Alam
- Lulwah Alkwai
- Mohamed Aturban
- Hussam Hallak
- Shawn Jones
- Mat Kelly
- Corren McCoy
- Louis Nguyen
- Alexander Nwala
- Nauman Siddique (MS)

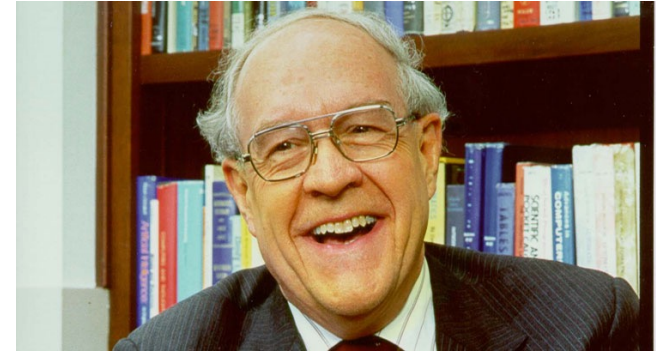
## Recent Alumni

- Yasmin AlNoamany
- Ahmed AlSum
- Grant Atkins (MS)
- John Berlin (MS)
- Justin Brunelle
- Chuck Cartledge
- Hung Do (MS)
- Maheedhar Gunnam (MS)
- Martin Klein
- Hany SalahEldeen
- Surbhi Shankar (MS)
- Erika Siregar (MS)
- Miranda Smith (MS)
- Plinio Vargas (MS)

@WebSciDL  
<http://ws-dl.cs.odu.edu/>  
<http://ws-dl.blogspot.com/>

# Computer scientists are toolsmiths

*The  
Computer  
Scientist  
as*  
**Toolsmith**



**F r e d e r i c k P . B r o o k s , J r .**

Frederick P. Brooks, Jr. 1996. The computer scientist as toolsmith II. Commun. ACM 39, 3 (March 1996), 61-68,  
<http://www.cs.unc.edu/~brooks/Toolsmith-CACM.pdf>



**UNC**  
DEPARTMENT OF  
COMPUTER SCIENCE

**We want to enable the  
personal use of web  
archives...**

# We want to enable the personal use of web archives... by academics and scholars

Liza Potts, ODU, Michigan State  
studying communication during disasters



A screenshot of a web browser displaying a BBC news article from the Internet Archive Wayback Machine. The browser address bar shows the URL: web.archive.org/web/20110723030954/http://www.bbc.com/. The Wayback Machine interface includes a search bar, a calendar showing the date July 23, 2011, and a "Go" button. The BBC logo and navigation menu (News, Sport, Weather, Travel, TV, Radio, More) are visible. The main content area features a "Top News story" section with the headline "'Eighty dead' in Norway shooting" and a sub-headline "Technology of Business" under a "Spotlight" section. The article text states: "At least 80 people died when a gunman opened fire at a Norwegian island youth camp, police say, hours after a deadly bomb attack on the capital, Oslo." A red banner at the bottom of the article reads "BREAKING NEWS". The "Spotlight" section includes a photo of a man in a brown jacket and a "CLOSED" sign in the background, with the text "Open all hours?" and "In our increasingly mobile and digital society, many retailers now see shops as an old-fashioned, inefficient way of doing business."

# They used screenshots to record news webpages and tweets

Name	Size	Kind	Date Modified
Disaster Oslo 2011	--	Folder	May 15, 2013 5:35 PM
BBC News - 'Tell the police to hurry. Peop...co.uk/news/world-europe-14287822.tiff	2.1 MB	TIFF image	Aug 6, 2011 7:08 PM
BBC News - Norway attacks/ The victims.pdf	456 KB	Portable Document Format (PDF)	Aug 2, 2011 2:13 PM
BBC News-Norway attacks-The victims_ht...co.uk/news/world-europe-14276074.tiff	3.8 MB	TIFF image	Aug 2, 2011 2:12 PM
Twitter_Anders_Lists.tiff	252 KB	TIFF image	Jul 23, 2011 12:00 AM
Twitter_Anders_Tweetstream.tiff	141 KB	TIFF image	Jul 23, 2011 12:00 AM
Twitter_Kjetil_trans.tiff	225 KB	TIFF image	Jul 23, 2011 12:00 AM
Twitter_Anders_Tweet.tiff	92 KB	TIFF image	Jul 22, 2011 11:59 PM
Twitter_Kjetil_stream.tiff	388 KB	TIFF image	Jul 22, 2011 11:59 PM
BBC - Homepage (20110722).jpg	837 KB	JPEG image	Jul 22, 2011 7:43 PM
CNN_Article_Update.jpg	5.1 MB	JPEG image	Jul 22, 2011 7:43 PM
CNN_Article.jpg	1.3 MB	JPEG image	Jul 22, 2011 7:43 PM
Giant blast hits government buildings in Oslo, Norway - CNN.com (20110722).jpg	1.3 MB	JPEG image	Jul 22, 2011 7:43 PM
Oslo.zip	21.8 MB	ZIP archive	Jul 22, 2011 7:42 PM
BBC - Homepage (20110722).jpg.zip	774 KB	ZIP archive	Jul 22, 2011 7:39 PM
Tragedie i Oslo og Utøya   Flickr - Photo Sharing! (20110722).tiff	806 KB	TIFF image	Jul 22, 2011 2:34 PM
Flickr_Brusse_Ok_5964328214.tiff	866 KB	TIFF image	Jul 22, 2011 2:33 PM
Bomb Explosion Oslo_1311357437671.png	301 KB	Portable Network Graphics image	Jul 22, 2011 1:58 PM
Bomb Explosion Oslo_1311357300058.png	1.1 MB	Portable Network Graphics image	Jul 22, 2011 1:57 PM
Screen shot 2011-07-22 at 1.51.47 PM.png	32 KB	Portable Network Graphics image	Jul 22, 2011 1:51 PM
MSM_Guardian_liveblog_TwitterUtøya.tiff	113 KB	TIFF image	Jul 22, 2011 1:48 PM
Screen shot 2011-07-22 at 1.46.54 PM.png	33 KB	Portable Network Graphics image	Jul 22, 2011 1:47 PM
MSM_Guardian_liveblog_621.tiff	655 KB	TIFF image	Jul 22, 2011 1:39 PM
Tweet_Fisher_TBvFB.tiff	77 KB	TIFF image	Jul 22, 2011 1:37 PM
BBC - Homepage Update.jpg	859 KB	JPEG image	Jul 22, 2011 1:36 PM
contentUSATODAY_update.jpg	2.2 MB	JPEG image	Jul 22, 2011 1:35 PM
USATODAY_Update.jpg	1 MB	JPEG image	Jul 22, 2011 1:34 PM
CNN_Update.jpg	1.9 MB	JPEG image	Jul 22, 2011 1:32 PM
Screen shot 2011-07-22 at 1.30.49 PM.png	25 KB	Portable Network Graphics image	Jul 22, 2011 1:30 PM
Screen shot 2011-07-22 at 1.25.24 PM.png	66 KB	Portable Network Graphics image	Jul 22, 2011 1:25 PM
USATODAY_Article.jpg	1.4 MB	JPEG image	Jul 22, 2011 11:42 AM
News, Travel, Weather, Entertainment, Sp...& World - USATODAY.com (20110722).jpg	1 MB	JPEG image	Jul 22, 2011 11:41 AM
BBC News - Oslo-Bomb blast near Norway prime minister's office (20110722).jpg	888 KB	JPEG image	Jul 22, 2011 11:41 AM
CNN_Page.jpg	1 MB	JPEG image	Jul 22, 2011 11:36 AM
Oslo_Paper.jpg	4.2 MB	JPEG image	Jul 22, 2011 11:34 AM

# We can find webpages for some filenames

Name	Size	Kind	Date Modified
Disaster Oslo 2011	--	Folder	May 15, 2013 5:35 PM
BBC News - 'Tell the police to hurry. Peop...co.uk/news/world-europe-14287822.tiff	2.1 MB	TIFF image	Aug 6, 2011 7:08 PM
BBC News - Norway attacks/ The victims.pdf	456 KB	Portable Document Format (PDF)	Aug 2, 2011 2:13 PM
BBC News - Norway attacks/ The victims - ...co.uk/news/world-europe-14276074.tiff	2.8 MB	TIFF image	Aug 2, 2011 2:13 PM



<http://www.bbc.com/news/world-europe-14287822>



<https://www.bbc.com/news/world-europe-14276074>

# But, it's difficult to manage metadata with just a filename

CNN_Article_Update.jpg	Jul 22, 2011 7:43 PM
CNN_Article.jpg	Jul 22, 2011 7:43 PM



# We want to enable the personal use of web archives... by academics and scholars

Alex Thurman and Pamela Graham



## Columbia course in Human Rights Information Technology

- evaluate online advocacy strategies over time
- explore the websites' degrees of interactivity
- observe the variety of ways groups frame and present issues online

# They want to view how groups' web presence changes over time

Alex Thurman and Pamela Graham



Internet Archive Wayback Machine screenshot showing search results for <http://amnesty.ca/>. The page title is "Human Rights Web Archive (Columbia University Libraries)".

Search results summary: Found 78 Captures between Mar 11, 2011 - Apr 18, 2018.

2011	2012	2013	2014	2015	2016	2017	2018
10 pages	11 pages	6 pages	7 pages	9 pages	16 pages	13 pages	6 pages
<a href="#">Mar 11, 2011</a> *	<a href="#">Mar 2, 2012</a> *	<a href="#">Mar 7, 2013</a> *	<a href="#">Jan 3, 2014</a> *	<a href="#">Jan 1, 2015</a> *	<a href="#">Jan 4, 2016</a> *	<a href="#">Jan 14, 2017</a> *	<a href="#">Jan 14, 2018</a> *
<a href="#">Mar 12, 2011</a> *	<a href="#">Mar 3, 2012</a> *	<a href="#">Mar 7, 2013</a> *	<a href="#">Jan 3, 2014</a> *	<a href="#">Apr 2, 2015</a> *	<a href="#">Jan 4, 2016</a> *	<a href="#">Jan 14, 2017</a> *	<a href="#">Jan 14, 2018</a> *
<a href="#">Apr 22, 2011</a> *	<a href="#">Mar 3, 2012</a>	<a href="#">Jun 11, 2013</a> *	<a href="#">Mar 27, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Jan 5, 2016</a> *	<a href="#">Jan 15, 2017</a> *	<a href="#">Jan 16, 2018</a> *
<a href="#">Jun 2, 2011</a> *	<a href="#">Jun 2, 2012</a> *	<a href="#">Jun 11, 2013</a> *	<a href="#">Mar 28, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 4, 2016</a> *	<a href="#">Apr 14, 2017</a> *	<a href="#">Apr 17, 2018</a> *
<a href="#">Jun 2, 2011</a> *	<a href="#">Jun 3, 2012</a> *	<a href="#">Oct 1, 2013</a> *	<a href="#">Jun 23, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 4, 2016</a> *	<a href="#">Apr 16, 2017</a> *	<a href="#">Apr 17, 2018</a> *
<a href="#">Jun 2, 2011</a>	<a href="#">Jun 3, 2012</a>	<a href="#">Oct 2, 2013</a> *	<a href="#">Jun 27, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 5, 2016</a> *	<a href="#">Jul 14, 2017</a> *	<a href="#">Apr 18, 2018</a> *
<a href="#">Sep 2, 2011</a> *	<a href="#">Jun 3, 2012</a>		<a href="#">Oct 1, 2014</a> *	<a href="#">Oct 1, 2015</a> *	<a href="#">Apr 6, 2016</a> *	<a href="#">Jul 14, 2017</a>	
<a href="#">Sep 3, 2011</a> *	<a href="#">Sep 2, 2012</a> *			<a href="#">Oct 2, 2015</a> *	<a href="#">May 11, 2016</a> *	<a href="#">Jul 16, 2017</a> *	
<a href="#">Dec 2, 2011</a> *	<a href="#">Sep 3, 2012</a> *			<a href="#">Dec 3, 2015</a> *	<a href="#">Jul 4, 2016</a> *	<a href="#">Oct 14, 2017</a> *	
<a href="#">Dec 3, 2011</a> *	<a href="#">Dec 2, 2012</a> *				<a href="#">Jul 4, 2016</a> *	<a href="#">Oct 14, 2017</a> *	
	<a href="#">Dec 2, 2012</a> *				<a href="#">Jul 5, 2016</a> *	<a href="#">Oct 15, 2017</a> *	
					<a href="#">Jul 6, 2016</a> *	<a href="#">Oct 15, 2017</a>	
					<a href="#">Jul 7, 2016</a> *	<a href="#">Oct 15, 2017</a> *	
					<a href="#">Oct 14, 2016</a> *		

[https://wayback.archive-it.org/1068/\\*/http://amnesty.ca/](https://wayback.archive-it.org/1068/*/http://amnesty.ca/)

# Visual layout changes are important

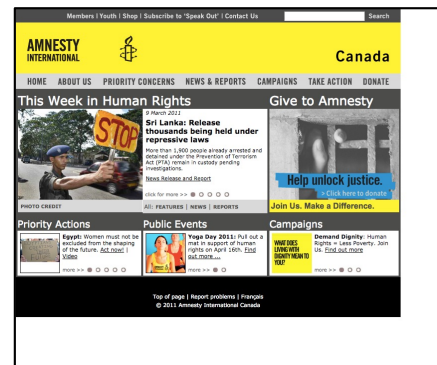
Alex Thurman and Pamela Graham



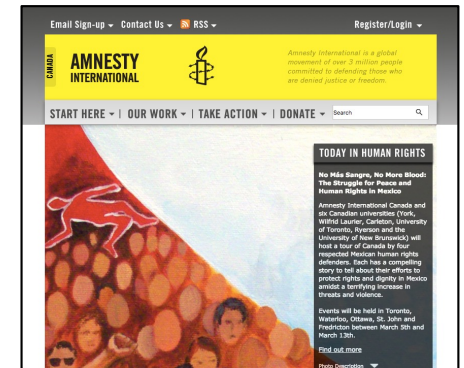
Internet Archive Wayback Machine screenshot showing the Human Rights Web Archive (Columbia University Libraries) search results for <http://amnesty.ca/>. The page displays 78 results found between Mar 11, 2011 and Apr 18, 2018.

Found 78 Captures between Mar 11, 2011 - Apr 18, 2018							
2011	2012	2013	2014	2015	2016	2017	2018
10 pages	11 pages	6 pages	7 pages	9 pages	16 pages	13 pages	6 pages
<a href="#">Mar 11, 2011</a> *	<a href="#">Mar 2, 2012</a> *	<a href="#">Mar 7, 2013</a> *	<a href="#">Jan 3, 2014</a> *	<a href="#">Jan 1, 2015</a> *	<a href="#">Jan 4, 2016</a> *	<a href="#">Jan 14, 2017</a> *	<a href="#">Jan 14, 2018</a> *
<a href="#">Mar 12, 2011</a> *	<a href="#">Mar 3, 2012</a> *	<a href="#">Mar 7, 2013</a> *	<a href="#">Jan 3, 2014</a> *	<a href="#">Apr 2, 2015</a> *	<a href="#">Jan 4, 2016</a> *	<a href="#">Jan 14, 2017</a> *	<a href="#">Jan 14, 2018</a> *
<a href="#">Apr 22, 2011</a> *	<a href="#">Mar 3, 2012</a> *	<a href="#">Jun 11, 2013</a> *	<a href="#">Mar 27, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Jan 5, 2016</a> *	<a href="#">Jan 15, 2017</a> *	<a href="#">Jan 16, 2018</a> *
<a href="#">Jun 2, 2011</a> *	<a href="#">Jun 2, 2012</a> *	<a href="#">Jun 11, 2013</a> *	<a href="#">Mar 28, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 4, 2016</a> *	<a href="#">Apr 14, 2017</a> *	<a href="#">Apr 17, 2018</a> *
<a href="#">Jun 2, 2011</a> *	<a href="#">Jun 3, 2012</a> *	<a href="#">Oct 1, 2013</a> *	<a href="#">Jun 23, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 4, 2016</a> *	<a href="#">Apr 16, 2017</a> *	<a href="#">Apr 17, 2018</a> *
<a href="#">Jun 2, 2011</a> *	<a href="#">Jun 3, 2012</a> *	<a href="#">Oct 2, 2013</a> *	<a href="#">Jun 27, 2014</a> *	<a href="#">Jul 1, 2015</a> *	<a href="#">Apr 5, 2016</a> *	<a href="#">Jul 14, 2017</a> *	<a href="#">Apr 18, 2018</a> *
<a href="#">Sep 2, 2011</a> *	<a href="#">Jun 3, 2012</a> *		<a href="#">Oct 1, 2014</a> *	<a href="#">Oct 1, 2015</a> *	<a href="#">Apr 6, 2016</a> *	<a href="#">Jul 14, 2017</a> *	
<a href="#">Sep 3, 2011</a> *	<a href="#">Sep 2, 2012</a> *			<a href="#">Oct 2, 2015</a> *	<a href="#">May 11, 2016</a> *	<a href="#">Jul 16, 2017</a> *	
<a href="#">Dec 2, 2011</a> *	<a href="#">Sep 3, 2012</a> *			<a href="#">Dec 3, 2015</a> *	<a href="#">Jul 4, 2016</a> *	<a href="#">Oct 14, 2017</a> *	
<a href="#">Dec 3, 2011</a> *	<a href="#">Dec 2, 2012</a> *				<a href="#">Jul 4, 2016</a> *	<a href="#">Oct 14, 2017</a> *	
	<a href="#">Dec 2, 2012</a> *				<a href="#">Jul 5, 2016</a> *	<a href="#">Oct 15, 2017</a> *	
					<a href="#">Jul 6, 2016</a> *	<a href="#">Oct 15, 2017</a> *	
					<a href="#">Jul 7, 2016</a> *	<a href="#">Oct 15, 2017</a> *	
					<a href="#">Oct 14, 2016</a> *		

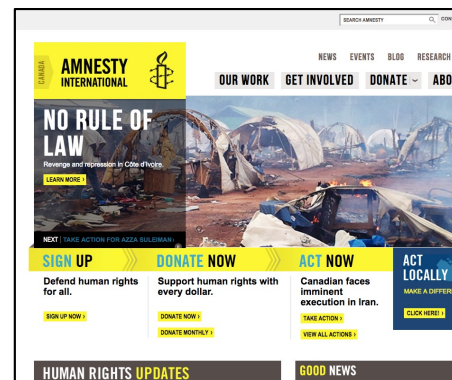
[https://wayback.archive-it.org/1068/\\*/http://amnesty.ca/](https://wayback.archive-it.org/1068/*/http://amnesty.ca/)



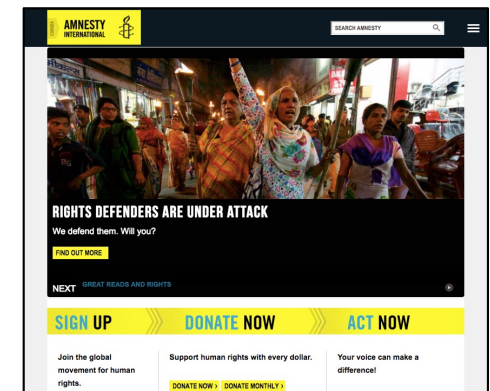
2011-03-11, 21:29:04



2012-03-02, 21:04:40



2013-03-07, 00:03:05



2018-01-14, 20:57:13

# We want to enable the personal use of web archives... by academics and scholars

Deborah Kempe THE FRICK COLLECTION



The screenshot shows a web browser window at archive-it.org/collections/4544. The page title is "Artists' Websites" and it was collected by the New York Art Resources Consortium (NYARC) in April 2014. The description states it contains websites of artists significant to the collections of the Brooklyn Museum, The Museum of Modern Art, and The Frick Collection. The subject is "Arts & Humanities, Society & Culture".

Under "Narrow Your Results", there is a search bar with the text "Enter search terms here" and buttons for "Search" and "Clear". Below the search bar, there are tabs for "Sites" and "Search Page Text".

The results section shows "Page 1 of 1 (22 Total Results)". The sorting options are "Title (A-Z)", "Title (Z-A)", "URL (A-Z)", and "URL (Z-A)".

The first result is "Robert Motherwell: The Dedalus Foundation" with URL <http://dedalusfoundation.org/motherwell>. A note says "No Captures were found for this URL".

The second result is "Isca Greenfield-Sanders" with URL <http://iscags.com/>. It was captured 6 times between Dec 5, 2014 and Dec 8, 2017.

The third result is "John Sanders" with URL <http://johnsandersart.com/>. It was captured 7 times between Dec 5, 2014 and Dec 8, 2017, and 3 videos were captured.

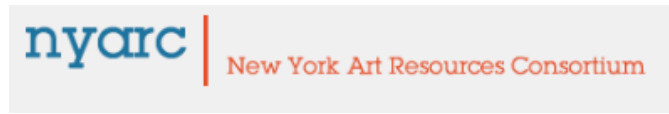
The fourth result is "Joop Sanders" with URL <http://joopsanders.com/>. It was captured 7 times between Dec 5, 2014 and Dec 8, 2017.

The fifth result is "Susan Hefuna" with URL <http://susanhefuna.com/>. It was captured 7 times between Dec 12, 2014 and Dec 8, 2017.

<https://archive-it.org/collections/4544>

# There's a need for visual browsing of collection of artists' websites

Deborah Kempe THE FRICK COLLECTION



archive-it.org/collections/4544

**nyarc**  
New York Art Resources Consortium

### Artists' Websites

Collected by: New York Art Resources Consortium (NYARC)

Archived since: Apr, 2014

Description: Websites of artists significant to the collections of the Brooklyn Museum, The Museum of Modern Art, and The Frick Collection.

Subject: Arts & Humanities, Society & Culture

**Narrow Your Results**

There are no further ways to narrow your results.

Enter search terms here

Sites

Page 1 of 1 (22 Total Results)

Sort By: [Title \(A-Z\)](#) | [Title \(Z-A\)](#) | [URL \(A-Z\)](#) | [URL \(Z-A\)](#)

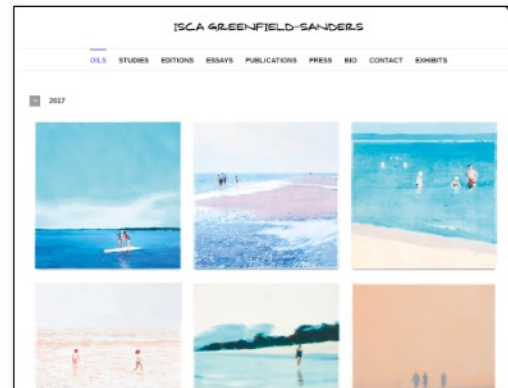
Title: Robert Motherwell: The Dedalus Foundation  
URL: <http://dedalusfoundation.org/motherwell>  
No Captures were found for this URL.

Title: Isca Greenfield-Sanders  
URL: <http://iscags.com/>  
Captured 6 times between Dec 5, 2014 and Dec 8, 2017

Title: John Sanders  
URL: <http://johnsandersart.com/>  
Captured 7 times between Dec 5, 2014 and Dec 8, 2017  
Videos: 3 Videos Captured

Title: Joop Sanders  
URL: <http://joopsanders.com/>  
Captured 7 times between Dec 5, 2014 and Dec 8, 2017

Title: Susan Hefuna  
URL: <http://susanhefuna.com/>  
Captured 7 times between Dec 12, 2014 and Dec 8, 2017



JOOP SANDERS

1940s 1950s 1960s Reviews / Essays Contact Information  
1970s 1980s - Present Biography

Grid of abstract paintings: a painting, a painting, a painting, a painting, a painting, a painting.

Death and Entrances III, 1950  
Oil on Linen

WORK

- ALL
- DRAWING
- PHOTOGRAPHY
- VIDEO
- INSTALLATION / SCULPTURE

EXHIBITIONS / PROJECTS  
BIOGRAPHY  
PUBLICATIONS  
LINKS  
CONTACT

SUSAN HEFUNA

Grid of artworks: a sculpture, a sculpture, a sculpture, a sculpture, a sculpture, a sculpture.

<https://archive-it.org/collections/4544>

# We want to enable the personal use of web archives... by journalists

## In 13 Headlines, the Drama of Election Night

<https://www.nytimes.com/2016/11/17/insider/in-13-headlines-the-drama-of-election-night.html>

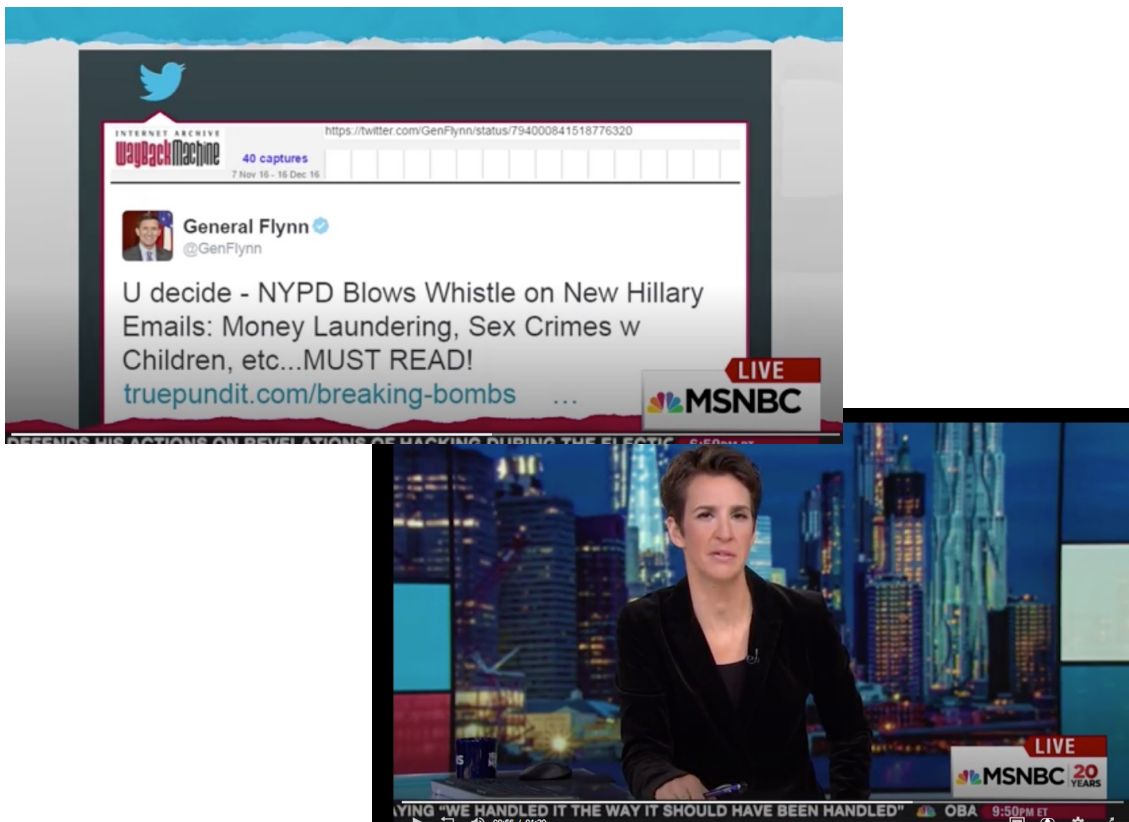
By Mark Bulik and Stephen Hiltner

Nov. 16, 2016

The screenshot shows the New York Times website interface during election night. At the top right, a black box displays "4:00 p.m.". The main headline is "In Final Hours of a Bitter Race, Voters Get Their Say". Below the headline is a large photograph of a line of people waiting outside a polling station. To the left of the photo are several article teasers, including "Result May Hinge on Changing Coalitions and Demographics" and "Battle for the Senate: 10 Races to Watch". To the right of the photo are more article teasers under the "OPINION" section, such as "What We're Seeing" and "A 'Mockery' of an Election". The website header includes the "The New York Times" logo, the date "Tuesday, November 8, 2016", and various navigation links.

similar to our Hurricane Katrina example: <https://www.slideshare.net/phonedude/why-careaboutthepast>

# Wayback has gone mainstream...



*Last Week Tonight, Mar 18, 2018*

"God bless you, Wayback Machine"  
- Rachel Maddow, Dec 16, 2016

# ... but what do people think the Wayback Machine is?

In the last week, though, additional posts on Reid's old blog were surfaced by Twitter user [@Jamie\\_Maz](#), who [also found](#) the posts about Crist. Reid's blog is no longer online, and she has said that she shuttered it in 2011. The offending posts were found via Internet Archive's [Wayback Machine](#), which archives website screenshots.

<https://www.politico.com/story/2018/04/25/joy-reid-anti-gay-posts-550213>



# ... but what do people think the Wayback Machine is?

In the last week, though, additional posts on Reid's old blog were surfaced by Twitter user [@Jamie\\_Maz](#), who [also found](#) the posts about Crist. Reid's blog is no longer online, and she has said that she shuttered it in 2011. The offending posts were found via Internet Archive's [Wayback Machine](#), which archives website screenshots.

<https://www.politico.com/story/2018/04/25/joy-reid-anti-gay-posts-550213>

<https://web.archive.org/web/20180115103952/https://auctionessistance.com/>

Pinedo's business, Auction Essistance, appears to have gone dark on the internet as of Friday, despite previously having Facebook and Twitter pages and a sizable amount of public customer reviews. But [cached versions of the business' homepage](#) and a [promotional Medium post](#) from 2015 show a business pitching itself as a trusted aid to webcrawlers who'd like privacy online.

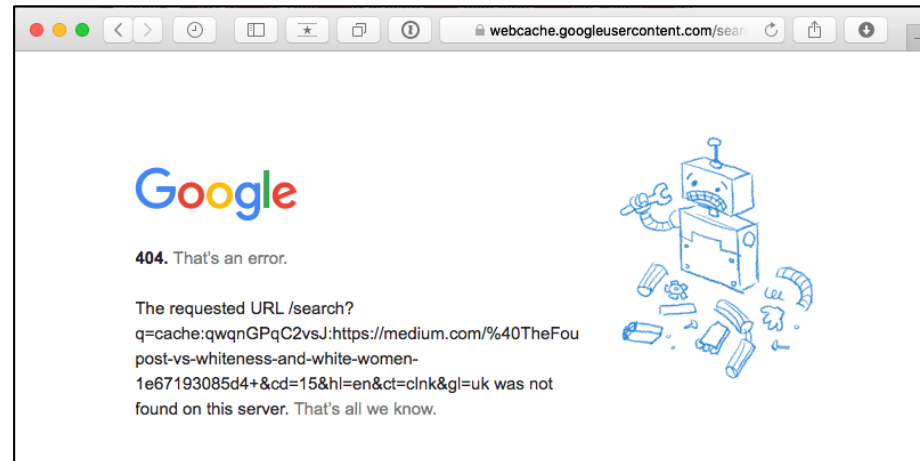
<https://www.cnn.com/2018/02/16/politics/richard-pinedo-guilty-plea/index.html>

# Caches are not archives

<http://www.wired.co.uk/article/russia-propaganda-online-blog-longform-medium-posts>

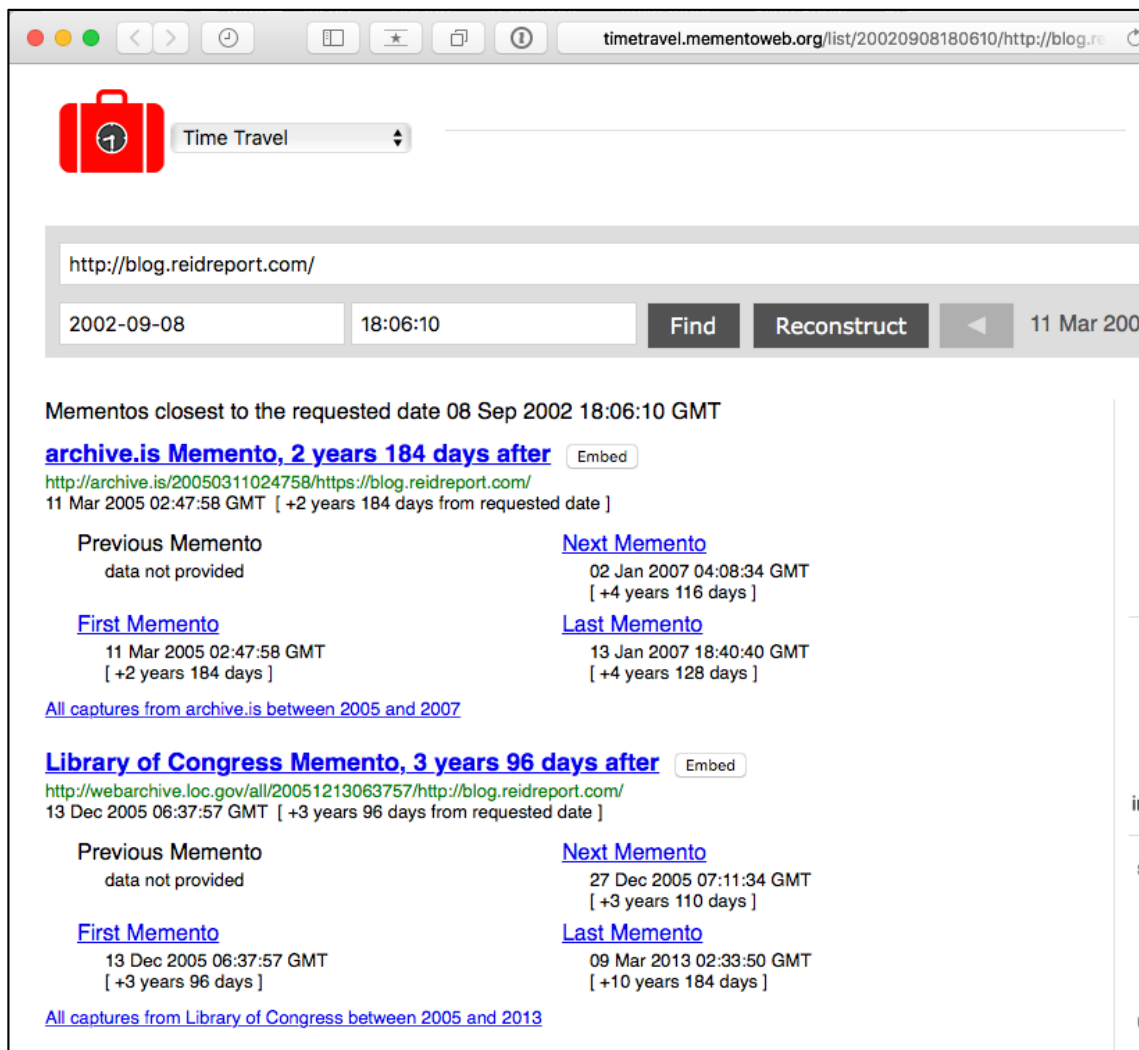
Twitter user @TheFoundingSon had more than 39,000 followers. The account, which went by the name John Davis, billed itself as a business-owning, Conservative, Christian family man. "Love my country and my family," its Twitter bio said. On Medium, @TheFoundingSon posted three blogs: [one on race](#), [one on science and politics](#), and [another on indoctrination at US colleges](#).

<https://webcache.googleusercontent.com/search?q=cache:qwqnGPqC2vsJ:https://medium.com/%40TheFoundingSon/huffington-post-vs-whiteness-and-white-women-1e67193085d4+&cd=15&hl=en&ct=clnk&gl=uk>



<http://ws-dl.blogspot.com/2018/01/2018-01-02-link-to-web-archives-not.html>

# And, there's more than just the Internet Archive



timetravel.mementoweb.org/list/20020908180610/http://blog.re

Time Travel

http://blog.reidreport.com/

2002-09-08 18:06:10 Find Reconstruct 11 Mar 200

Mementos closest to the requested date 08 Sep 2002 18:06:10 GMT

[archive.is Memento, 2 years 184 days after](#) Embed  
<http://archive.is/20050311024758/https://blog.reidreport.com/>  
11 Mar 2005 02:47:58 GMT [ +2 years 184 days from requested date ]

<a href="#">Previous Memento</a> data not provided	<a href="#">Next Memento</a> 02 Jan 2007 04:08:34 GMT [ +4 years 116 days ]
<a href="#">First Memento</a> 11 Mar 2005 02:47:58 GMT [ +2 years 184 days ]	<a href="#">Last Memento</a> 13 Jan 2007 18:40:40 GMT [ +4 years 128 days ]

[All captures from archive.is between 2005 and 2007](#)

[Library of Congress Memento, 3 years 96 days after](#) Embed  
<http://webarchive.loc.gov/all/20051213063757/http://blog.reidreport.com/>  
13 Dec 2005 06:37:57 GMT [ +3 years 96 days from requested date ]

<a href="#">Previous Memento</a> data not provided	<a href="#">Next Memento</a> 27 Dec 2005 07:11:34 GMT [ +3 years 110 days ]
<a href="#">First Memento</a> 13 Dec 2005 06:37:57 GMT [ +3 years 96 days ]	<a href="#">Last Memento</a> 09 Mar 2013 02:33:50 GMT [ +10 years 184 days ]

[All captures from Library of Congress between 2005 and 2013](#)



install Memento for Chrome

enable web time travel

 **MEDIAWIKI**

install Memento for MediaWiki


say **no** to "404 Not Found"

 **Robust Links**

use links that refuse to die

<http://timetravel.mementoweb.org/list/20020908180610/http://blog.reidreport.com/>

# Some folks knows this

 <http://archive.is/SKYbp>

[One of the posts](#) also criticizes Rachel Maddow, who is now the most-watched host on MSNBC, as being “at the left-most end of the political spectrum” when it comes to gay rights. (Ms. Reid hosts a weekend morning show on MSNBC, but at the time she kept her blog active neither she nor Ms. Maddow had shows on the network.)

<https://www.nytimes.com/2018/04/24/business/media/joy-reid-homophobic-blog-posts.html>

# Some folks knows this

 <http://archive.is/SKYbp>

One of the posts also criticizes Rachel Maddow, who is now the most-watched host on MSNBC, as being “at the left-most end of the political spectrum” when it comes to gay rights. (Ms. Reid hosts a weekend morning show on MSNBC, but at the time she kept her blog active neither she nor Ms. Maddow had shows on the network.)

<https://www.nytimes.com/2018/04/24/business/media/joy-reid-homophobic-blog-posts.html>

A Library of Congress archive of the site shows that the "lengthy" entry contains only two sentences of text actually written by the post's author; the rest is a quote.

The Library of Congress archive reviewed by CNNMoney -- which the Library says is created using a local installation of the Wayback Machine -- contains the disputed posts and lists them as having been archived on January 12, 2006. The documents provided by MSNBC to CNNMoney do not contain a letter to the Library of Congress regarding its archive.

<http://money.cnn.com/2018/04/25/media/joy-reid-msnbc-host-wayback-machine/index.html>

# We advocate submitting pages to multiple archives

**nex**    
@botherder Follow

Tip: if you really want to archive webpages before someone gets to delete them, always take a screenshot. Last year I spent days archiving all I could from the website of FlexiSpy stalkerware, and @internetarchive somehow decided months later to wipe it all and exclude the site.



4:22 AM - 12 May 2018

[https://twitter.com/phonedude\\_mln/status/998948823845261312](https://twitter.com/phonedude_mln/status/998948823845261312)

**Michael L. Nelson**   
@phonedude\_mln Following

Replying to @botherder @internetarchive

FYI: a few pages exist in other archives:

[wayback.archive-it.org/all/2009080301 ...](http://wayback.archive-it.org/all/2009080301...)  
[wayback.archive-it.org/all/2010061901 ...](http://wayback.archive-it.org/all/2010061901...)  
[collection.europarchive.org/nli/2011031103 ...](http://collection.europarchive.org/nli/2011031103...)  
...  
[archive.is/20130508132946 ...](http://archive.is/20130508132946...)  
[wayback.archive-it.org/all/2013091006 ...](http://wayback.archive-it.org/all/2013091006...)

in the future, push to \*multiple\* archives:

[ws-dl.blogspot.com/2017/02/2017-0...](http://ws-dl.blogspot.com/2017/02/2017-0...)

screenshots are nice, but not "evidence"



```
$ archivenow --ia www.cnn.com
https://web.archive.org/web/20170220182349/http://www.cnn.com

Captured by The Internet Archive and Archive.is

$ archivenow --ia --is www.cnn.com
https://web.archive.org/web/20170220185038/http://www.cnn.com
http://archive.is/s5kvj

Captured by all archives

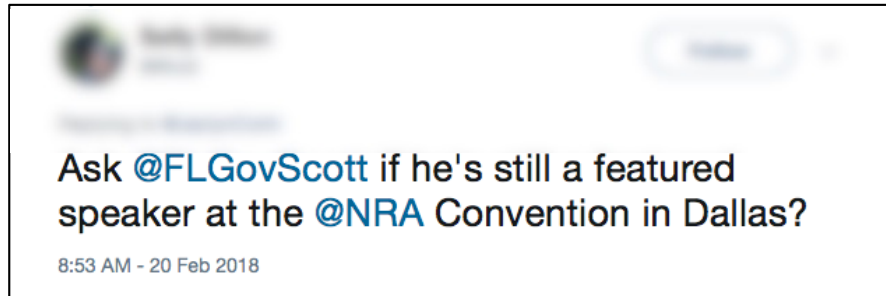
$ archivenow --all www.cnn.com --cc_api_key $Your-API-Key
https://perma.cc/8V8H-3FAH
https://web.archive.org/web/20170220185656/http://www.cnn.com

2017-02-22: Archive Now (archivenow): A Python Library to Integrate On-De...
Examples: Archive Now (archivenow) CLI A small part of my research is to ensure
that certain web pages are preserved in public web archi...

ws-dl.blogspot.com
```

11:28 AM - 22 May 2018

# We want to enable the personal use of web archives... by the general public



# Web archives to the rescue!

Ask [@FLGovScott](#) if he's still a featured speaker at the [@NRA](#) Convention in Dallas?  
8:53 AM - 20 Feb 2018


Looks like his picture is now deleted from [@NRA](#)'s confirmed speakers page? I knew I should have taken a screenshot before [@FLGovScott](#) weaseled out!  
12:03 PM - 20 Feb 2018

The screenshot shows a Twitter thread from Brian E. Griffin (@brian3354) replying to @Essie\_82, @8fuld, and 3 others. The tweet text is: "That's why web archives are important. [web.archive.org/web/2018021917...](https://web.archive.org/web/2018021917...)" The tweet is dated 5:46 PM - 20 Feb 2018 and has 4 retweets and 28 likes. Below the tweet are two replies: "This is awesome" and "You, Sir, are today's hero!". An inset image shows a web browser displaying the NRA website's "ANNUAL MEETINGS" page. The page lists four confirmed speakers: Rep. Richard Hudson, Tomi Lahren (Political Commentator), Governor Rick Scott, and Attorney General Adam Laxalt. The browser's address bar shows the URL: <https://www.nraam.org/special-events/2018-events/nra-ta-leadership-forum/>. The browser's date is set to Feb 19, 2018.

<https://twitter.com/brian3354/status/966081774194511874>



# Is it really that important to archive instead of just taking a screenshot?

**Imani Gandy**  [@AngryBlackLady](#) Follow

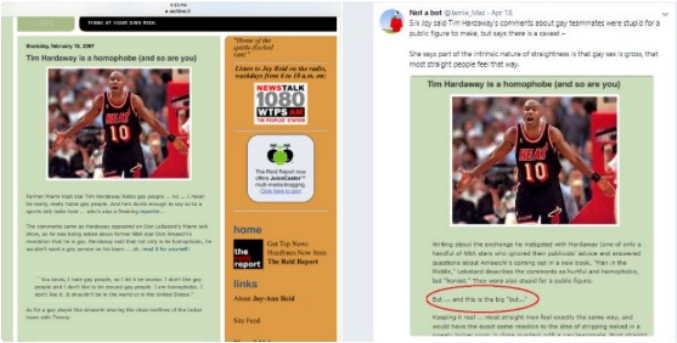
Replying to [@WalkerBragman](#) [@MSNBC](#)

I just screencapped your link. That's the first image. The second image is what Maz posted.

The images are different. Do you see?

Look at the first line of image 1 versus the first line of image 2.

Care to explain? I'm seriously asking. Perhaps I'm missing something.



8:58 PM - 27 Apr 2018


<https://twitter.com/AngryBlackLady/status/990032514080108544>

**Michael L. Nelson** [@phonedude\\_mln](#) Following

Replying to [@mpersandy](#) [@PhoenixWomanMN](#)

the screen shots are from [@Jamie\\_Maz](#), so I can't say for sure. here are two things to consider:

1. I suspect the Hardaway image is added to the top of the text for attention; note that all of the text in the screen shot is in fact at: [archive.is/20070218173925...](https://archive.is/20070218173925...) but 50-60% down.



11:28 PM - 27 Apr 2018

[https://twitter.com/phonedude\\_mln/status/990070331737100288](https://twitter.com/phonedude_mln/status/990070331737100288)

# We should be doing both

**Conspirador Norteño** @conspirator0 · May 26

This account changed names sometime during that four-year gap. Based on this snapshot from the Internet Archive, @veteranhnak was @louiebaur in 2012. Curiously, there's a @louiebaur account now with tweets from 2011, indicating that at least two accounts were renamed.

Google TV's second wind: New devices coming from LG, Samsung, Vizio and Sony dvr/13wFF

**Veteran Hank** @veteranhank · 5 Jan 2012  
Mobile ad network Millennial Media files for iPod dvr/13wFF

**Veteran Hank** @veteranhank · 5 Jan 2012  
Replying to @Gitterlee  
@Gitterlee I sure am did it work?

**Veteran Hank** @veteranhank · 5 Jan 2012  
Replying to @nanospartak  
@nanospartak LOL totally

**Veteran Hank** @veteranhank · 5 Jan 2012  
Replying to @BobWarren  
@BobWarren You to man!

**Veteran Hank** @veteranhank · 5 Jan 2012  
I haven't heard Casey Anthony in awhile

**louiebaur** Louie Baer  
Google TV's second wind: New devices coming from Samsung, Vizio and Sony dvr/13wFF

**louiebaur** Louie Baer  
Mobile ad network Millennial Media files for iPod dvr/13wFF

**louiebaur** Louie Baer  
@Gitterlee I sure am did it work?

**louiebaur** Louie Baer  
@nanospartak LOL totally

**louiebaur** Louie Baer  
@BobWarren You to man!

**BobWarren** Bob Warren  
The Dirty Word LinkedIn Game: tryurl.com/20efaz

**louiebaur** Louie Baer  
I haven't heard Casey Anthony in awhile

2 9 40

**Conspirador Norteño** @conspirator0 **Follow**

(Archive link the screenshot in the previous tweet was taken from)  
[web.archive.org/web/2012010617...](https://web.archive.org/web/2012010617...)

4:33 PM - 26 May 2018

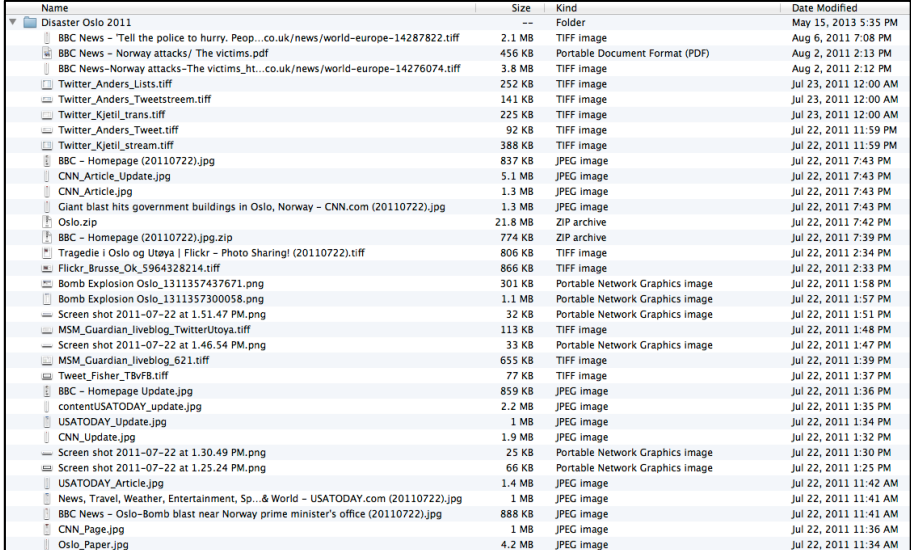
4 Retweets 27 Likes

1 4 27

<https://twitter.com/conspirator0/status/1000475042017366017>

**What have we been doing  
to make this easier?**

# We wanted to help people create and access local archives



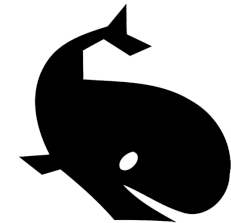
Name	Size	Kind	Date Modified
Disaster Oslo 2011	--	Folder	May 15, 2013 5:35 PM
BBC News - Tell the police to hurry. Peop...co.uk/news/world-europe-14287822.tiff	2.1 MB	TIFF image	Aug 6, 2011 7:08 PM
BBC News - Norway attacks/ The victims.pdf	456 KB	Portable Document Format (PDF)	Aug 2, 2011 2:13 PM
BBC News - Norway attacks- The victims_ht...co.uk/news/world-europe-14276074.tiff	3.8 MB	TIFF image	Aug 2, 2011 2:12 PM
Twitter_Anders_Lists.tiff	252 KB	TIFF image	Jul 23, 2011 12:00 AM
Twitter_Anders_Tweetstream.tiff	141 KB	TIFF image	Jul 23, 2011 12:00 AM
Twitter_Kjetil_trans.tiff	225 KB	TIFF image	Jul 22, 2011 12:00 AM
Twitter_Anders_Tweet.tiff	92 KB	TIFF image	Jul 22, 2011 11:59 PM
Twitter_Kjetil_stream.tiff	388 KB	TIFF image	Jul 22, 2011 11:59 PM
BBC - Homepage (20110722).jpg	837 KB	JPEG image	Jul 22, 2011 7:43 PM
CNN_Article_Update.jpg	5.1 MB	JPEG image	Jul 22, 2011 7:43 PM
CNN_Article.jpg	1.3 MB	JPEG image	Jul 22, 2011 7:43 PM
Giant blast hits government buildings in Oslo, Norway - CNN.com (20110722).jpg	1.3 MB	JPEG image	Jul 22, 2011 7:43 PM
Oslo.zip	21.8 MB	ZIP archive	Jul 22, 2011 7:42 PM
BBC - Homepage (20110722).jpg.zip	774 KB	ZIP archive	Jul 22, 2011 7:39 PM
Tragedie i Oslo og Utaya   Flickr - Photo Sharing! (20110722).tiff	806 KB	TIFF image	Jul 22, 2011 2:34 PM
Flickr_Brusse_Ok_5964328214.tiff	866 KB	TIFF image	Jul 22, 2011 2:33 PM
Bomb Explosion Oslo_1311357437671.png	301 KB	Portable Network Graphics image	Jul 22, 2011 1:58 PM
Bomb Explosion Oslo_1311357300058.png	1.1 MB	Portable Network Graphics image	Jul 22, 2011 1:57 PM
Screen shot 2011-07-22 at 1.51.47 PM.png	32 KB	Portable Network Graphics image	Jul 22, 2011 1:51 PM
MSM_Guardian_liveblog_TwitterUtoya.tiff	113 KB	TIFF image	Jul 22, 2011 1:48 PM
Screen shot 2011-07-22 at 1.46.54 PM.png	33 KB	Portable Network Graphics image	Jul 22, 2011 1:47 PM
MSM_Guardian_liveblog_621.tiff	655 KB	TIFF image	Jul 22, 2011 1:39 PM
Tweet_Fisher_TBvE8.tiff	77 KB	TIFF image	Jul 22, 2011 1:37 PM
BBC - Homepage Update.jpg	859 KB	JPEG image	Jul 22, 2011 1:36 PM
contentUSATODAY_update.jpg	2.2 MB	JPEG image	Jul 22, 2011 1:35 PM
USATODAY_Update.jpg	1 MB	JPEG image	Jul 22, 2011 1:34 PM
CNN_Update.jpg	1.9 MB	JPEG image	Jul 22, 2011 1:32 PM
Screen shot 2011-07-22 at 1.30.49 PM.png	25 KB	Portable Network Graphics image	Jul 22, 2011 1:30 PM
Screen shot 2011-07-22 at 1.25.24 PM.png	66 KB	Portable Network Graphics image	Jul 22, 2011 1:25 PM
USATODAY_Article.jpg	1.4 MB	JPEG image	Jul 22, 2011 11:42 AM
News, Travel, Weather, Entertainment, Sp... & World - USATODAY.com (20110722).jpg	1 MB	JPEG image	Jul 22, 2011 11:41 AM
BBC News - Oslo-Bomb blast near Norway prime minister's office (20110722).jpg	888 KB	JPEG image	Jul 22, 2011 11:41 AM
CNN_Page.jpg	1 MB	JPEG image	Jul 22, 2011 11:36 AM
Oslo_Paper.jpg	4.2 MB	JPEG image	Jul 22, 2011 11:34 AM

# We wanted to help people create and access local archives

- *WARCreate* – Google Chrome extension



- *WAIL* – user-friendly Heritrix and OpenWayback



- *WAIL-Electron* – adds browser-based crawling, pywb

# WARCreate (2012)



Google Chrome extension

Create local WARC file of currently viewed webpage

<http://warcreate.com>

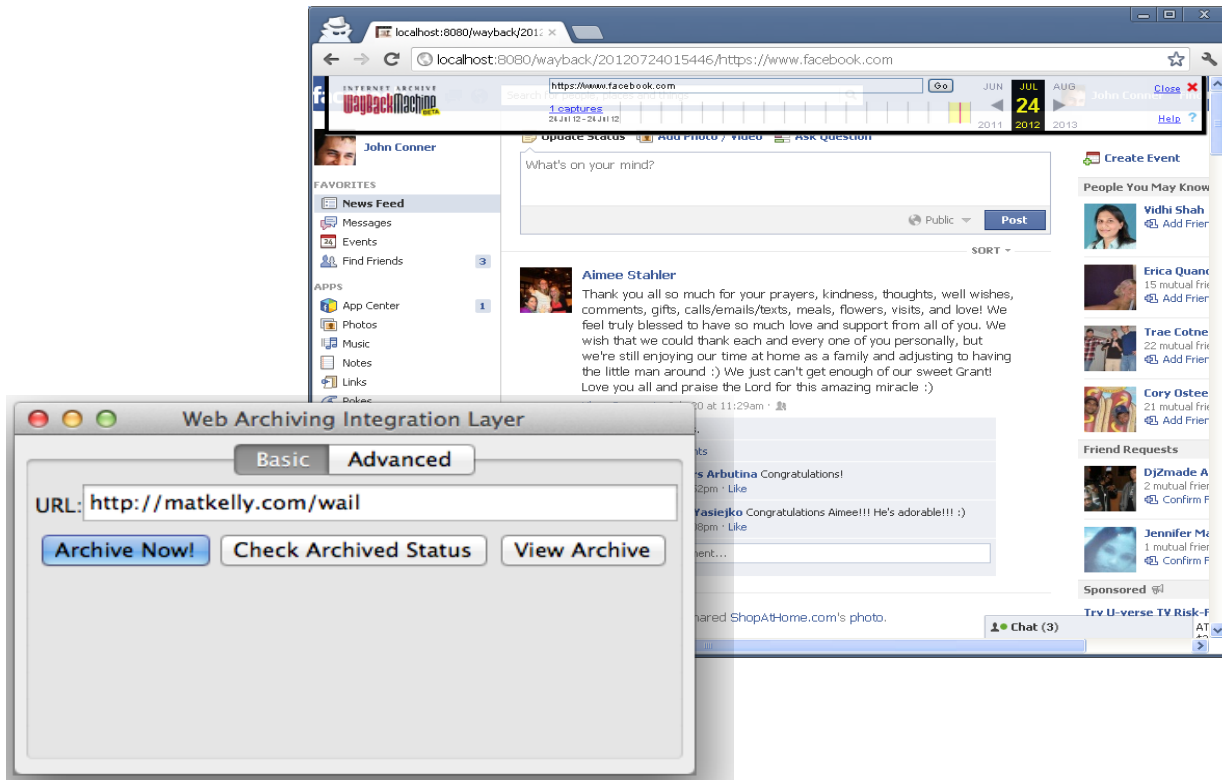
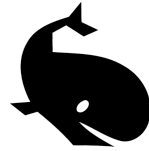
**Mat Kelly** and Michele C. Weigle, "WARCreate - Create Wayback-Consumable WARC Files from Any Webpage", JCDL 2012 demo.

<http://ws-dl.blogspot.com/2013/07/2013-07-10-warcreate-and-wail-warc.html>



"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher", 2013-2017, HD-51670-13 and HK-50181-14

# WAIL (2013)



Stand-alone application

Easy install of Heritrix,  
OpenWayback

Replay local WARC's created  
with WARCreate

<http://machawk1.github.io/wail/>

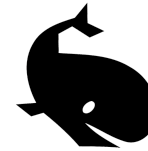
**Mat Kelly**, Michael L. Nelson and Michele C. Weigle, "Making Enterprise-Level Archive Tools Accessible for Personal Web Archiving Using XAMPP," Poster and demo at Personal Digital Archiving, 2013.

<http://ws-dl.blogspot.com/2016/06/2016-06-03-lipstick-or-ham-next-steps.html>



"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher", 2013-2017, HD-51670-13 and HK-50181-14

# WAIL-Electron (2017)



WAIL

Collections > default > Add Seed FROM FILESYSTEM

Seed to add:  
http://example.com

Archive Configuration  
Page + Same domain links

CHECK SEED

HTTP 200 OK  
External Links: 1

ADD AND ARCHIVE NOW CANCEL

Update of original WAIL

Adds headless Chrome-based crawling

OpenWayback -> pywb

<https://github.com/N0taN3rd/wail>

**John Berlin**, Mat Kelly, Michael L. Nelson and Michele C. Weigle, "WAIL: Collection-Based Personal Web Archiving," JCDL 2017, poster.

<http://ws-dl.blogspot.com/2017/02/2017-02-13-electric-wails-and-ham.html>

<http://ws-dl.blogspot.com/2017/07/2017-07-24-replacing-heritrix-with.html>



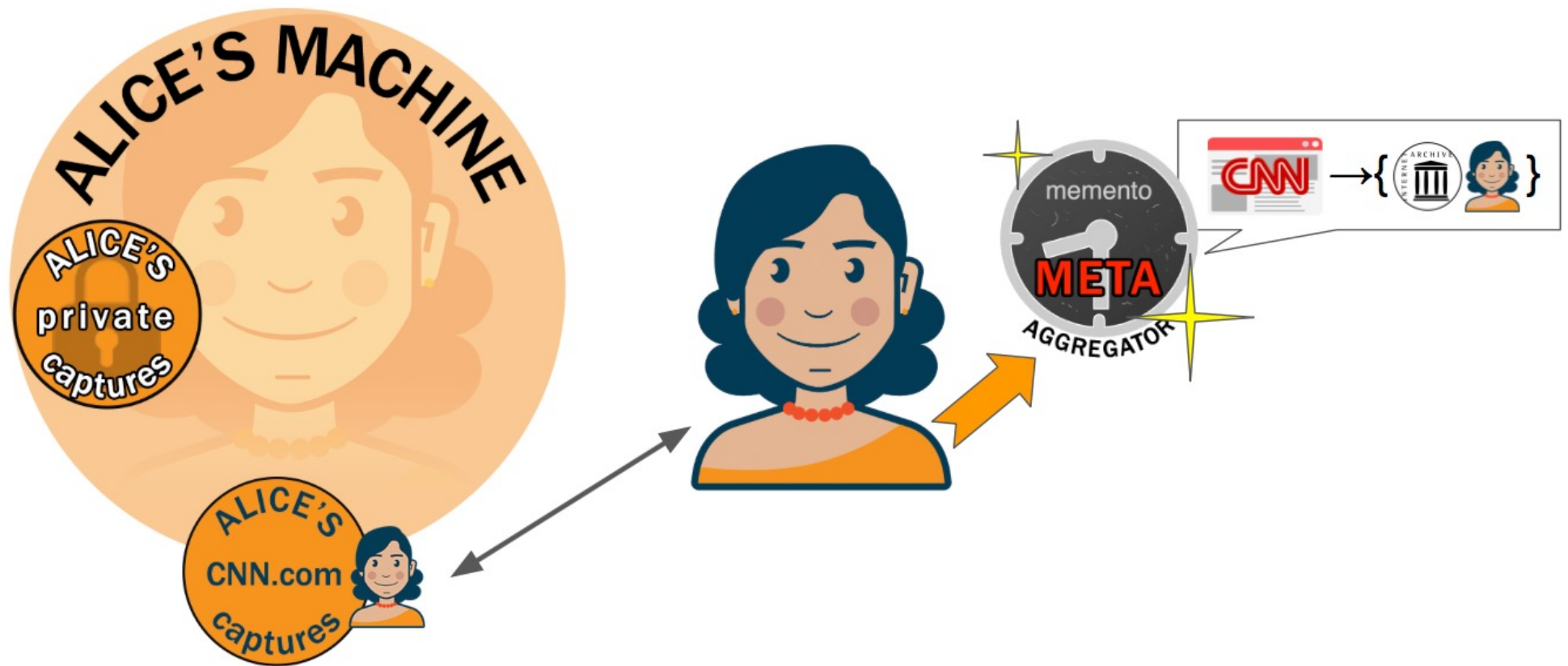
"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher",  
2013-2017, HD-51670-13 and HK-50181-14



# What did we learn from this?

- We need additional Memento support for private web archives
- Capturing complex webpages is hard

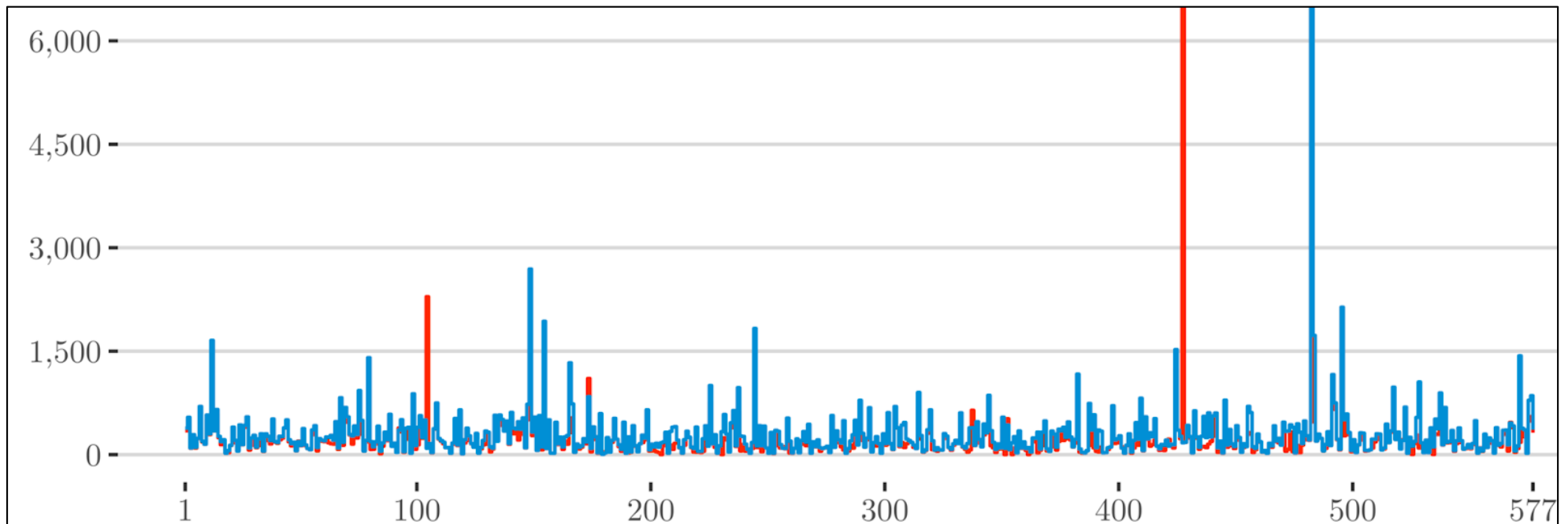
# A Memento Meta Aggregator can aggregate public and private archives (2018)



Mat Kelly, Michael L. Nelson, and Michele C. Weigle, "A Framework for Aggregating Private and Public Web Archives", JCDL 2018

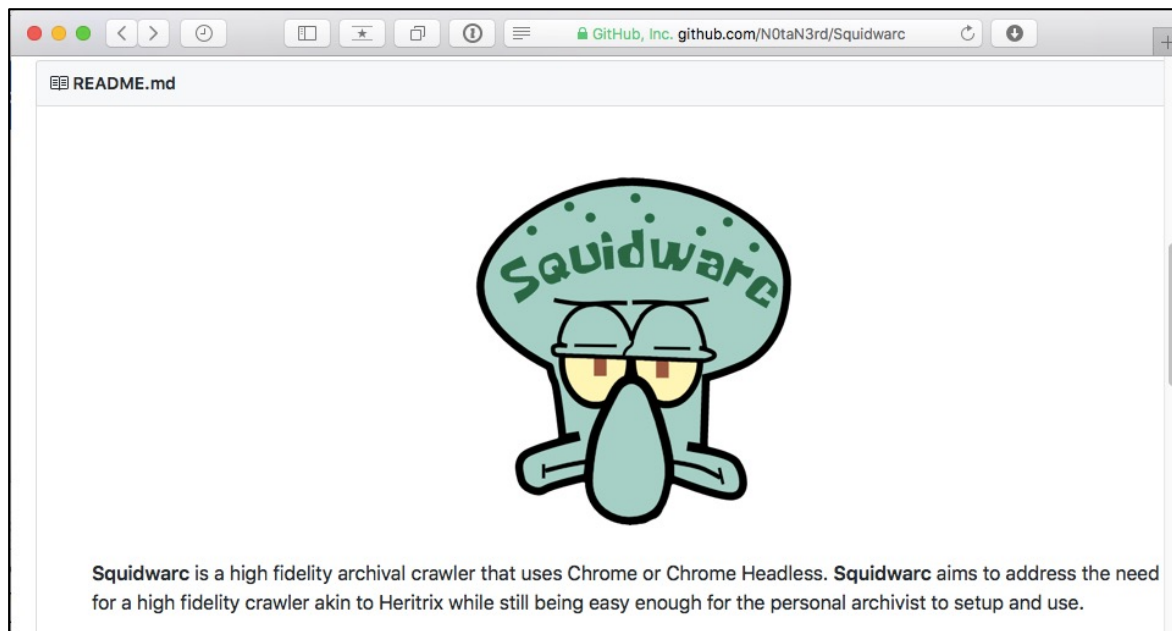
# Today's webpages are super complex

number of network requests per page



**John Berlin**, "To Relive The Web: A Framework for the Transformation and Archival Replay of Web Pages," ODU Master's Thesis, 2018.

# Squidwarc enables high-fidelity browser-based archiving (2017)



High fidelity archival crawler

node.js based

Uses Chrome or Chrome Headless

<https://github.com/N0taN3rd/Squidwarc>

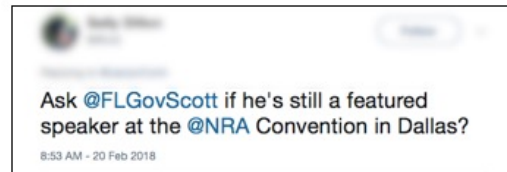
**John Berlin**, "2017-07-24: Replacing Heritrix with Chrome in WAIL, and the release of node-warc, node-cdxj, and Squidwarc"

<http://ws-dl.blogspot.com/2017/07/2017-07-24-replacing-heritrix-with.html>



"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher", 2013-2017, HD-51670-13 and HK-50181-14

# We wanted to help people submit webpages to public archives



# We wanted to help people submit webpages to public archives

- *Mink* – Google Chrome extension



- *#icanhazmemento* – Twitter bot



- *ArchiveNow* – Python module, Docker container, local web service

# Mink (2014)



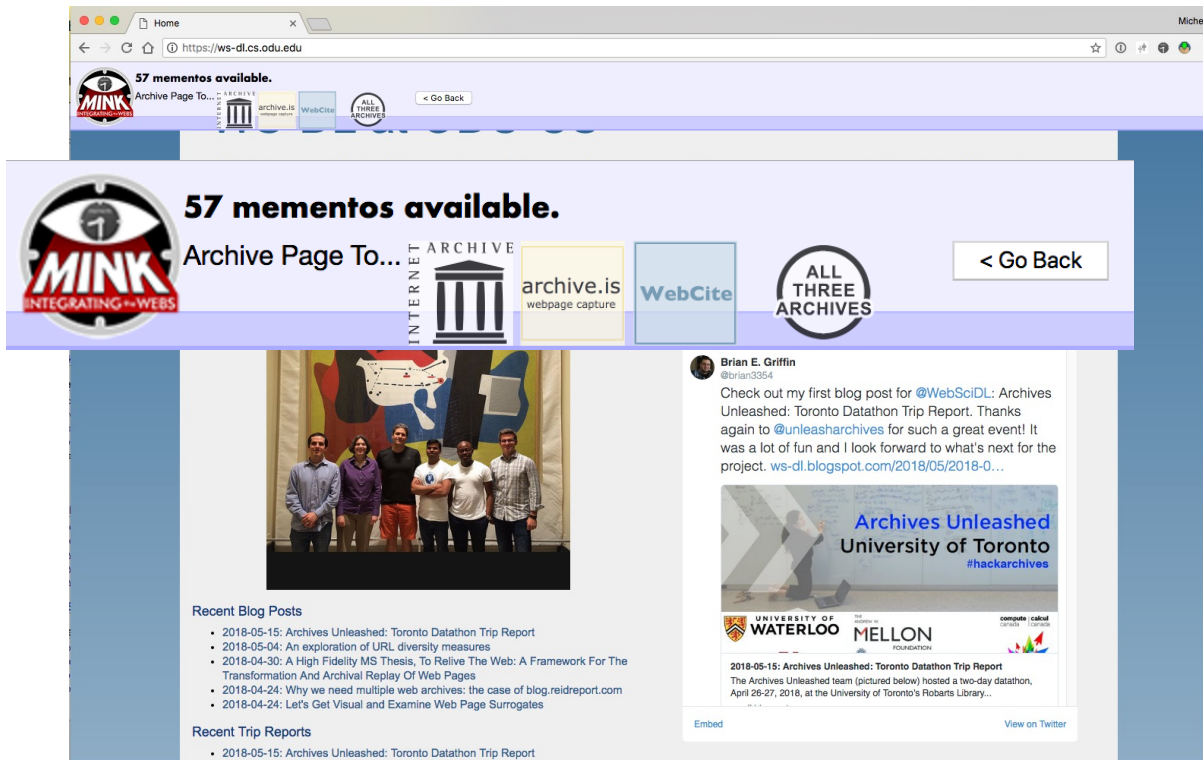
Google Chrome extension

Submit currently viewed webpage to public archives

Access mementos from public archives of currently viewed webpage

Inspired by LANL's Memento for Chrome, <http://ws-dl.blogspot.com/2013/10/2013-10-14-right-click-to-past-memento.html>

<https://github.com/machawk1/Mink>



**Mat Kelly**, Michael L. Nelson and Michele C. Weigle, "Mink: Integrating the Live and Archived Web Viewing Experience Using Web Browsers and Memento," JCDL 2014, poster.

<http://ws-dl.blogspot.com/2014/10/2014-10-03-integrating-live-and.html>



"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher", 2014-2017, HK-50181-14

# Mink (2014)



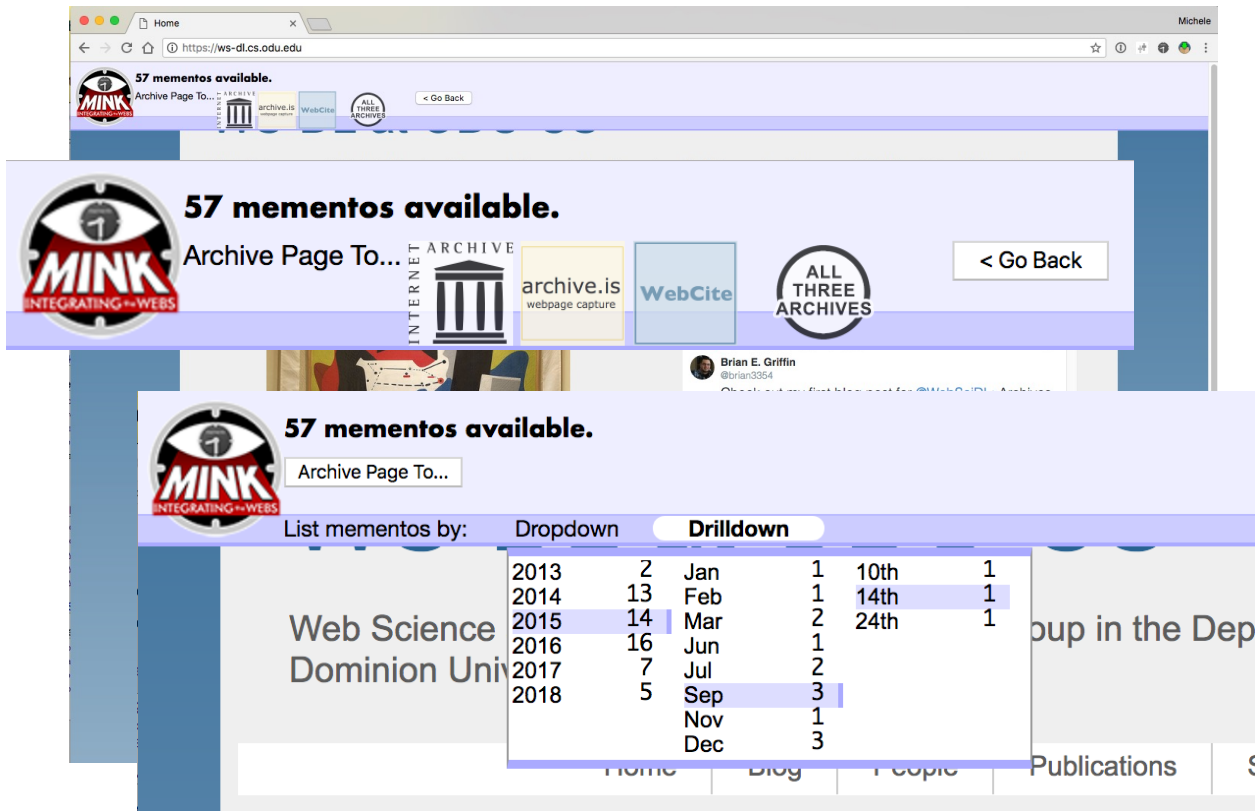
Google Chrome extension

Submit currently viewed webpage to public archives

Access mementos from public archives of currently viewed webpage

Inspired by LANL's Memento for Chrome, <http://ws-dl.blogspot.com/2013/10/2013-10-14-right-click-to-past-memento.html>

<https://github.com/machawk1/Mink>



**Mat Kelly**, Michael L. Nelson and Michele C. Weigle, "Mink: Integrating the Live and Archived Web Viewing Experience Using Web Browsers and Memento," JCDL 2014, poster.

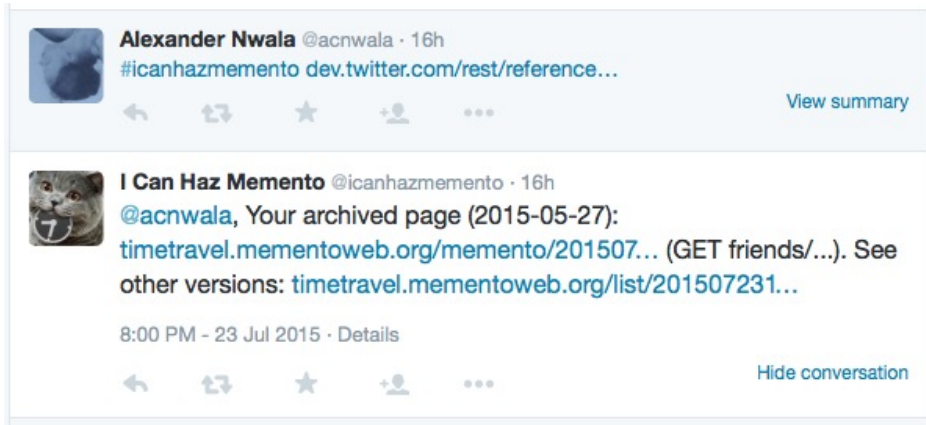
<http://ws-dl.blogspot.com/2014/10/2014-10-03-integrating-live-and.html>



"Archive What I See Now: Bringing Institutional Web Archiving Tools to the Individual Researcher", 2014-2017, HK-50181-14



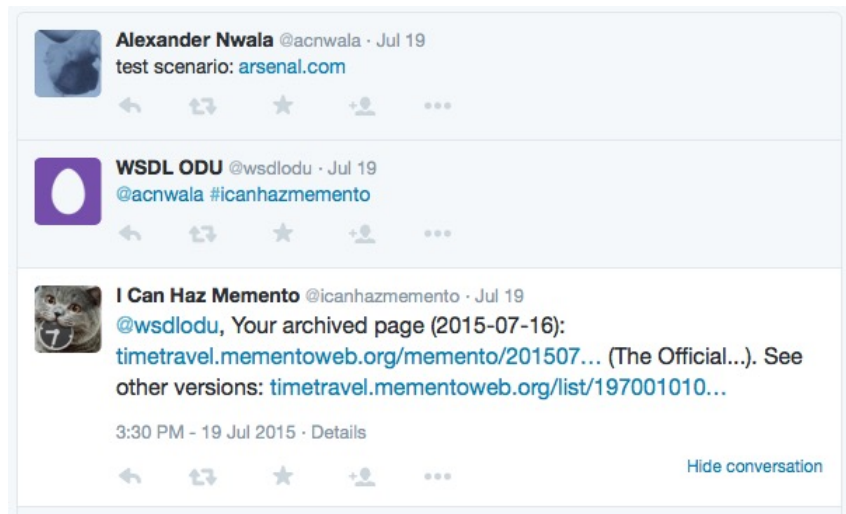
# #icanhazmemento (2015)



Twitter bot

Include #icanhazmemento in a tweet with a URL

Bot replies with a link to the memento of the page closest to the time of the tweet



If page not archived, bot submits URL to multiple public archives, replies with a link to the memento in Time Travel

<https://github.com/anwala/icanhazmemento>

Alexander Nwala, "2015-07-22: I Can Haz Memento,"  
<http://ws-dl.blogspot.com/2015/07/2015-07-22-i-can-haz-memento.html>

# ArchiveNow (2017)

```
$ archivenow --all --cc_api_key=6af4f...df72c www.example.com
https://perma.cc/SMC3-LRQT
https://web.archive.org/web/20180527081945/http://www.example.com
http://archive.is/IM0Az
http://www.webcitation.org/6zitVQdbw
www.example.com.warc

$ archivenow --ia --is www.example.com
https://web.archive.org/web/20180527084957/http://www.example.com
http://archive.is/5MKea

$ archivenow --warc=myArchive --agent=wget www.example.com
myArchive.warc

$ archivenow --server --port=8085
* Running on http://0.0.0.0:8085/ (Press CTRL&C to quit)
```

Python module, Docker container

Submit URI to multiple archives

Generate local WARCs for private archives

<https://github.com/oduwsdl/archivenow>

**Mohamed Aturban**, Mat Kelly, Sawood Alam, John Berlin, Michael L. Nelson and Michele C. Weigle, "ArchiveNow: Simplified, Extensible, Multi-Archive Preservation," JCDL 2018, poster.

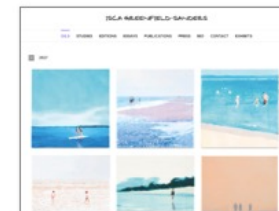
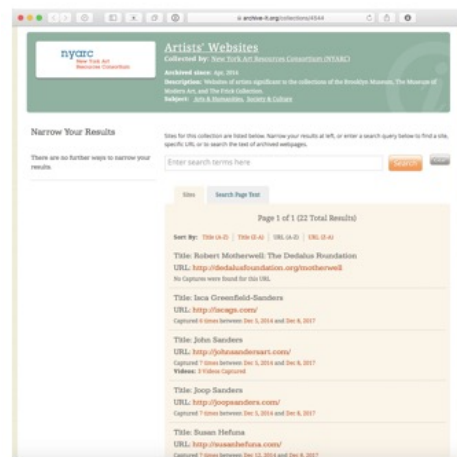
<http://ws-dl.blogspot.com/2017/02/2017-02-22-archive-now-archivenow.html>

THE ANDREW W.  
**MELLON** "Towards a Web-Centric Approach for Capturing the Scholarly Record", 2016-2019  
FOUNDATION

# What did we learn from this?

- People want tools to help them submit to public archives
- Browser extensions are cool, but don't have much uptake
- *more on this later...*

# We wanted to help people summarize their archives

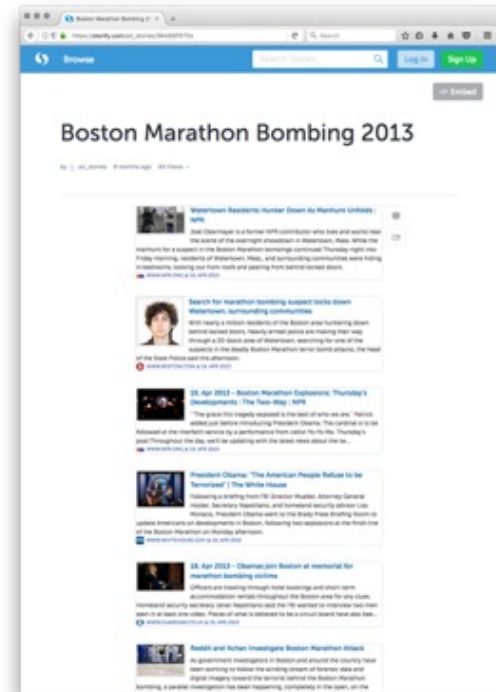
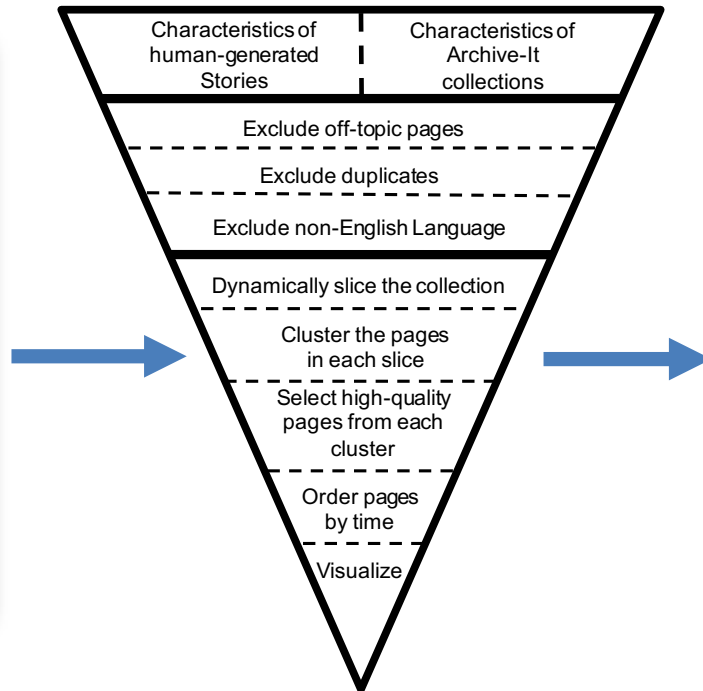
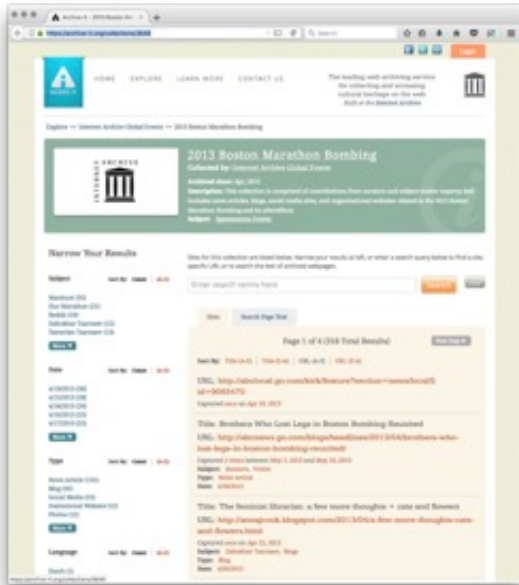


# We wanted to help people summarize their archives

- *Dark and Stormy Archives (DSA)* – Archive-It + Storify
- *MementoEmbed* – web service
- *#whatdiditlooklike* – Twitter bot
- *Alsummarization* – algorithm and web service
- *TimeMap Visualization, tmvis* – node.js-based web service of alsummarization



# "Dark and Stormy" Archives (2016)



**Yasmin AlNoamany**, Michele C. Weigle, and Michael L. Nelson, "Generating Stories From Archived Collections," ACM WebSci 2017.

<http://ws-dl.blogspot.com/2016/09/2016-09-20-promising-scene-at-end-of.html>

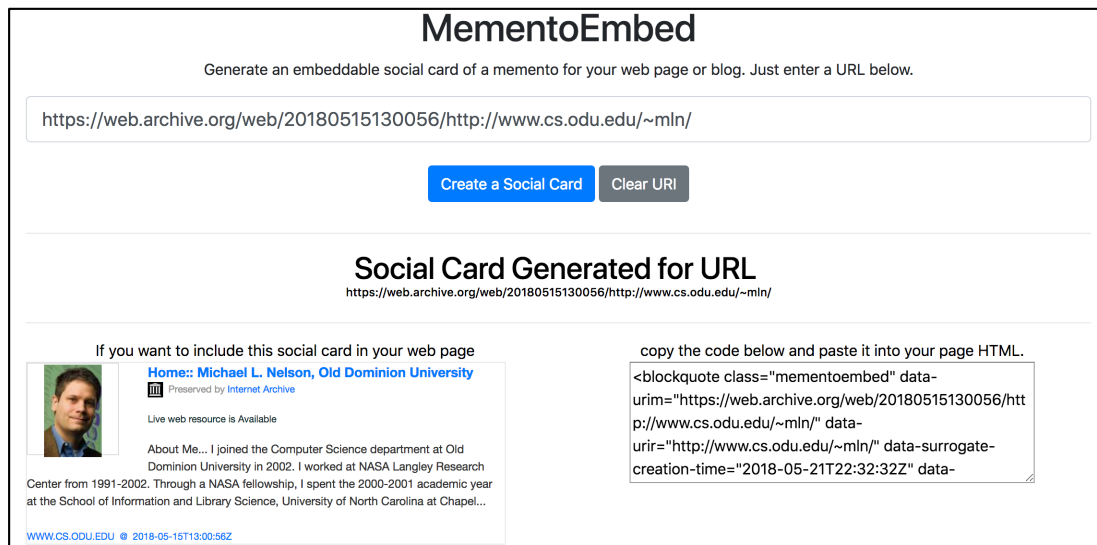
**Shawn Jones**, "Improving Collection Understanding in Web Archives," JCDL Doctoral Consortium, 2018.

<http://ws-dl.blogspot.com/2017/12/2017-12-14-storify-will-be-gone-soon-so.html>



"Combining Social Media Storytelling With Web Archives", 2015-2019, IMLS National Leadership Grant

# MementoEmbed (2018)




**MementoEmbed**  
Generate an embeddable social card of a memento for your web page or blog. Just enter a URL below.

[Create a Social Card](#) [Clear URI](#)

**Social Card Generated for URL**  
<https://web.archive.org/web/20180515130056/http://www.cs.odu.edu/~mln/>

If you want to include this social card in your web page

 **Home: Michael L. Nelson, Old Dominion University**  
Preserved by Internet Archive  
Live web resource is Available

About Me... I joined the Computer Science department at Old Dominion University in 2002. I worked at NASA Langley Research Center from 1991-2002. Through a NASA fellowship, I spent the 2000-2001 academic year at the School of Information and Library Science, University of North Carolina at Chapel...

copy the code below and paste it into your page HTML.

```
<blockquote class="mementoembed" data-urim="https://web.archive.org/web/20180515130056/http://www.cs.odu.edu/~mln/" data-urir="http://www.cs.odu.edu/~mln/" data-surrogate-creation-time="2018-05-21T22:32:32Z" data-
```

WWW.CS.ODU.EDU © 2018-05-15T13:00:56Z

Python module, Docker container

Submit URI-M

Returns an archive-aware social card, with HTML embed code

<http://mementoembed.ws-dl.cs.odu.edu/>

<https://github.com/oduwsdl/MementoEmbed>

**Shawn Jones**, "Improving Collection Understanding in Web Archives," JCDL Doctoral Consortium, 2018.

<http://ws-dl.blogspot.com/2018/04/2018-04-24-lets-get-visual-and-examine.html>



“Combining Social Media Storytelling With Web Archives”, 2015-2019, IMLS National Leadership Grant

# MementoEmbed (2018)

The screenshot shows the MementoEmbed web interface. At the top, it says "MementoEmbed" and "Generate an embeddable social card of a memento for your web page or blog. Just enter a URL below." There is a text input field containing the URL "https://web.archive.org/web/20180515130056/http://www.cs.odu.edu/~mln/". Below the input field are two buttons: "Create a Social Card" and "Clear URI".

Below the buttons, it says "Social Card Generated for URL" followed by the same URL. Underneath, there are instructions: "If you want to include this social card in your web page" and "copy the code below and paste it into your page HTML." There are two small preview windows. The left one shows a profile picture and the text "Center from 1991-2002 at the School of Information and Library Science, University of North Carolina at Chapel Hill". The right one shows the HTML code: "<blockquote class='mementoembed' data-...".

The main social card generated is for "Home: Michael L. Nelson, Old Dominion University". It includes a profile picture, a small icon, and the text "Live web resource is Available". Below that is a bio: "About Me... I joined the Computer Science department at Old Dominion University in 2002. I worked at NASA Langley Research Center from 1991-2002. Through a NASA fellowship, I spent the 2000-2001 academic year at the School of Information and Library Science, University of North Carolina at Chapel Hill...". At the bottom of the card is a footer: "CS WWW.CS.ODU.EDU @ 2018-05-15T13:00:56Z".

Python module, Docker container

Submit URI-M

Returns an archive-aware social card, with HTML embed code

<http://mementoembed.ws-dl.cs.odu.edu/>

<https://github.com/oduwsdl/MementoEmbed>

**Shawn Jones**, "Improving Collection Understanding in Web Archives," JCDL Doctoral Consortium, 2018.

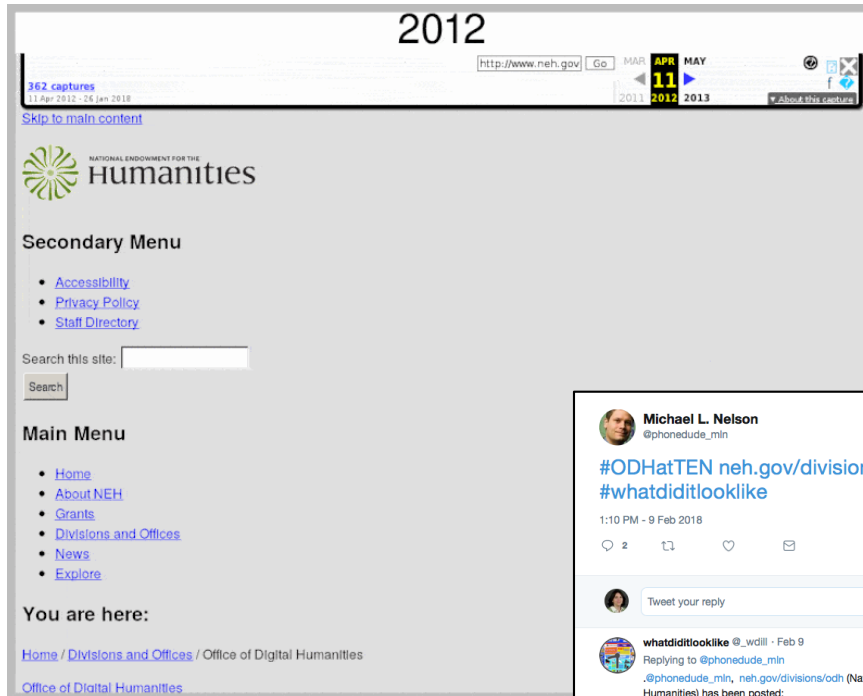
<http://ws-dl.blogspot.com/2018/04/2018-04-24-lets-get-visual-and-examine.html>



"Combining Social Media Storytelling With Web Archives", 2015-2019, IMLS National Leadership Grant



# #whatdiditlooklike (2015)



Twitter bot

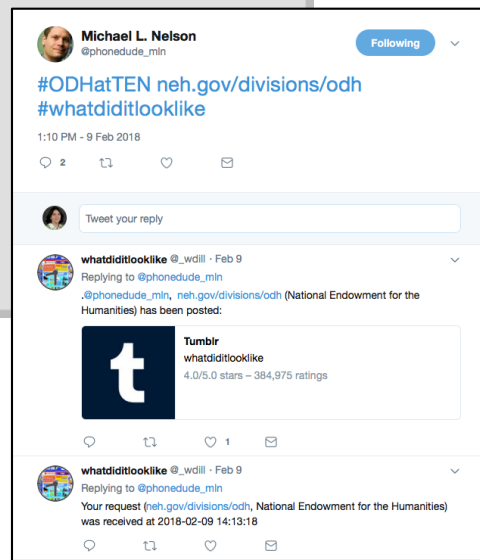
Include #whatdiditlooklike in a tweet with a URL

Bot generates animated GIF of first memento of each year

Bot replies with a link to entry in Tumblr

Tumblr:  
<http://whatdiditlooklike.mementoweb.org/>

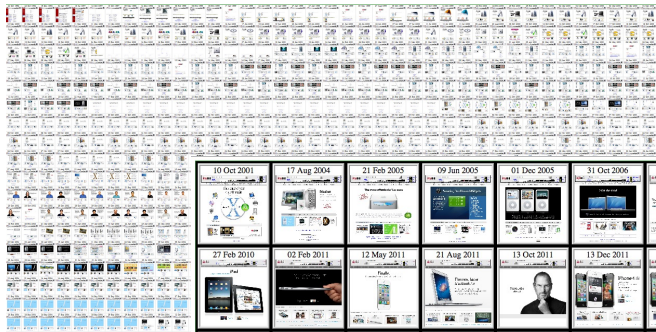
Source:  
<https://github.com/anwala/wdill>



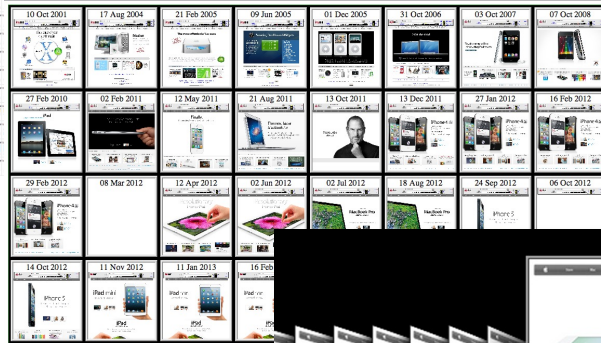
Alexander Nwala, "2015-02-05: What Did It Look Like?,"

<http://ws-dl.blogspot.com/2015/01/2015-02-05-what-did-it-look-like.html>

# Alsummarization (2014)



700 thumbnails



32 sampled thumbnails



CoverFlow view

Summarize TimeMap

Compare SimHash of HTML, not images

Hamming distance threshold of 4 characters

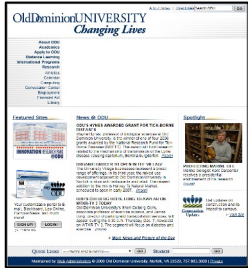
<https://github.com/machawk1/ArchiveThumbnails>

**Ahmed Alsum** and Michael L. Nelson, "Thumbnail Summarization Techniques for Web Archives," ECIR 2014.

**Mat Kelly**, Michael L. Nelson, and Michele C. Weigle, "Visualizing Digital Collections of Web Archives," Web Archiving Collaboration, 2015, <http://ws-dl.blogspot.com/2015/06/2015-06-09-web-archiving-collaboration.html>

# Choosing mementos based on SimHash

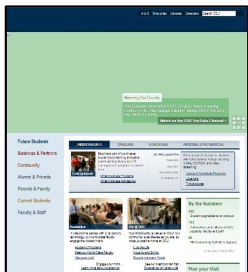
M1



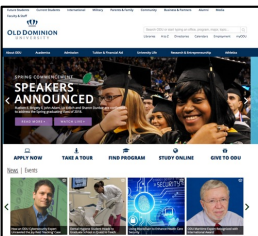
M2



M3

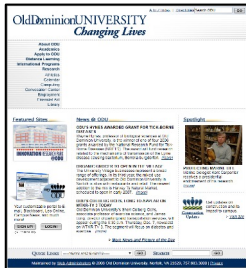


M4



# Choosing mementos based on SimHash

M1



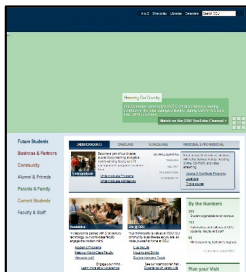
8c27981eaed151cfa645ad823932eac6

M2



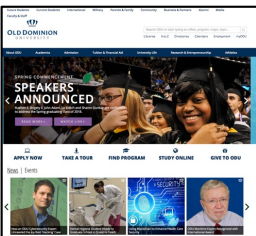
8c27981eaad951cf8645ad823932eac6

M3



fa3799170258494b9443b9be3977a84e

M4



5a1534161357da6b827ab98037db2640

# Choosing mementos based on SimHash

M1



8c27981eaed151cfa645ad823932eac6



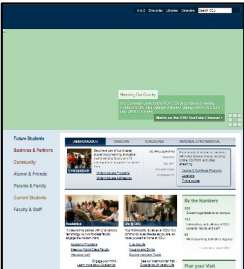
M1

M2



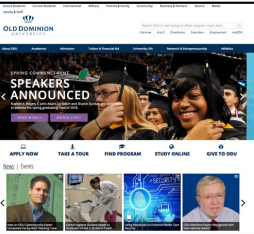
8c27981eaad951cf8645ad823932eac6

M3



fa3799170258494b9443b9be3977a84e

M4



5a1534161357da6b827ab98037db2640

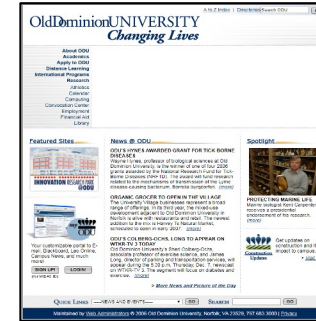
# Choosing mementos based on SimHash

M1



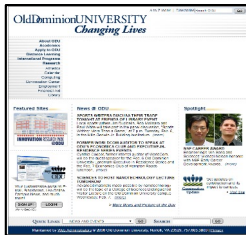
8c27981eaed151cfa645ad823932eac6

basis



M1

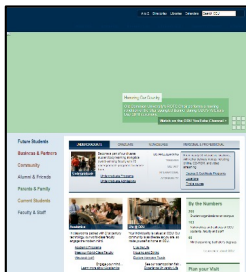
M2



8c27981ead951cf8645ad823932eac6

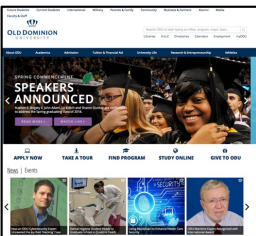
Hamming distance (M1, M2) < 4  
*reject M2*

M3



fa3799170258494b9443b9be3977a84e

M4



5a1534161357da6b827ab98037db2640

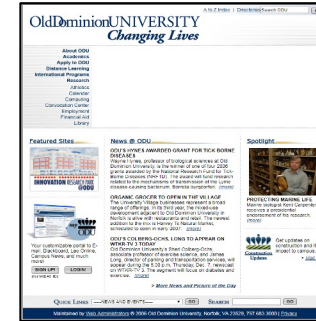
# Choosing mementos based on SimHash

M1



8c27981eaed151cfa645ad823932eac6

basis



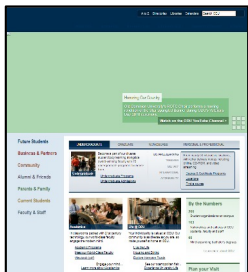
M1

M2



8c27981eaad951cf8645ad823932eac6

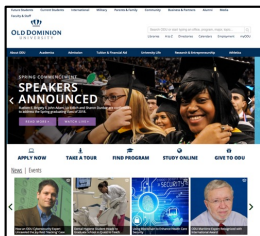
M3



fa3799170258494b9443b9be3977a84e

Hamming distance (M1, M3) > 4  
*select M3*

M4



5a1534161357da6b827ab98037db2640

# Choosing mementos based on SimHash

M1



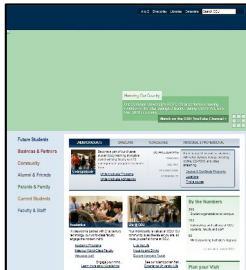
8c27981eaed151cfa645ad823932eac6

M2



8c27981eaad951cf8645ad823932eac6

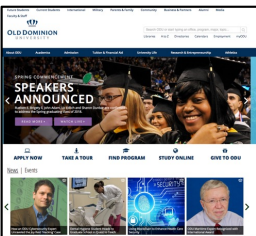
M3



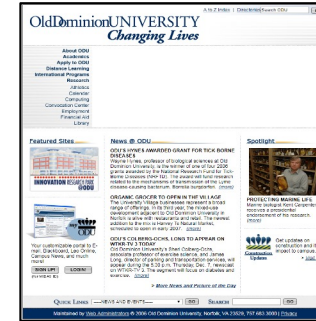
fa3799170258494b9443b9be3977a84e

basis

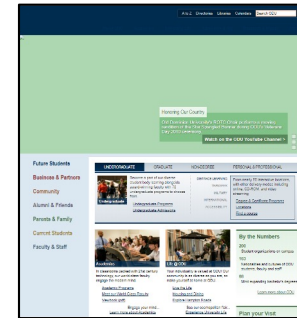
M4



5a1534161357da6b827ab98037db2640



M1



M3

Hamming distance (M3, M4) > 4  
select M4



# Choosing mementos based on SimHash

M1



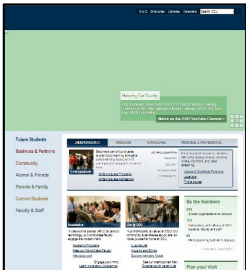
8c27981eaed151cfa645ad823932eac6

M2



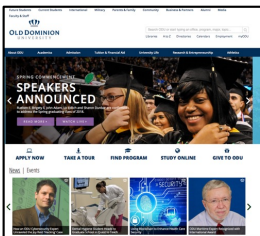
8c27981ead951cf8645ad823932eac6

M3

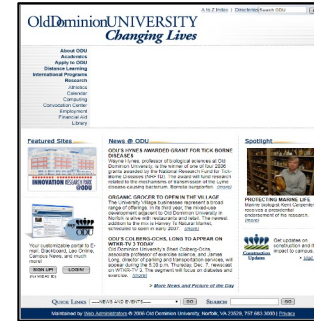


fa3799170258494b9443b9be3977a84e

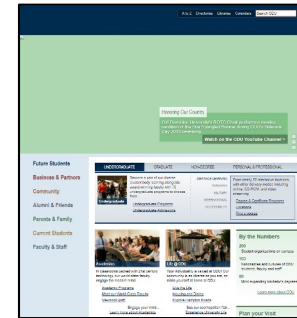
M4



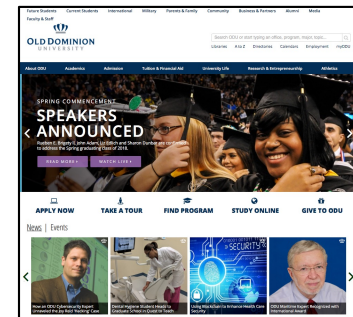
5a1534161357da6b827ab98037db2640



M1



M3



M4

# TimeMap Visualization, tmvis (2017)

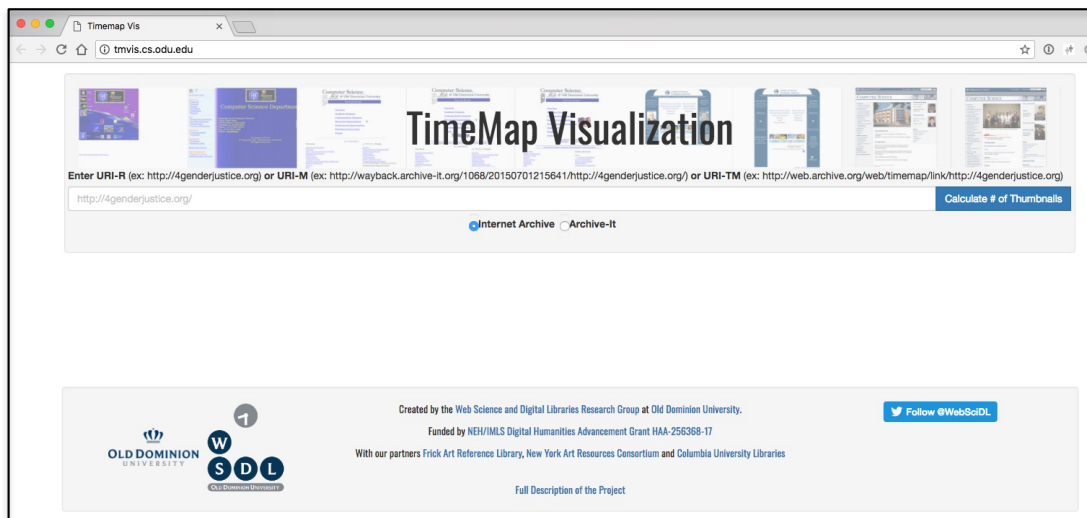
Web service

Takes URI-R or URI-T

Performs *Alsummarization* and produces grid view, image slider view, and timeline view

Will produce embeddable version, Wayback extension

<https://github.com/oduwsdl/tmvis>



**Surbhi Shankar**, "Visualizing Thumbnails Of Archived Web Pages", ODU MS Project, 2017

**Maheedhar Gunnam**, "How I Changed Over Time: A webservice to summarize TimeMaps based on SimHashed HTML content", ODU MS Project, 2018

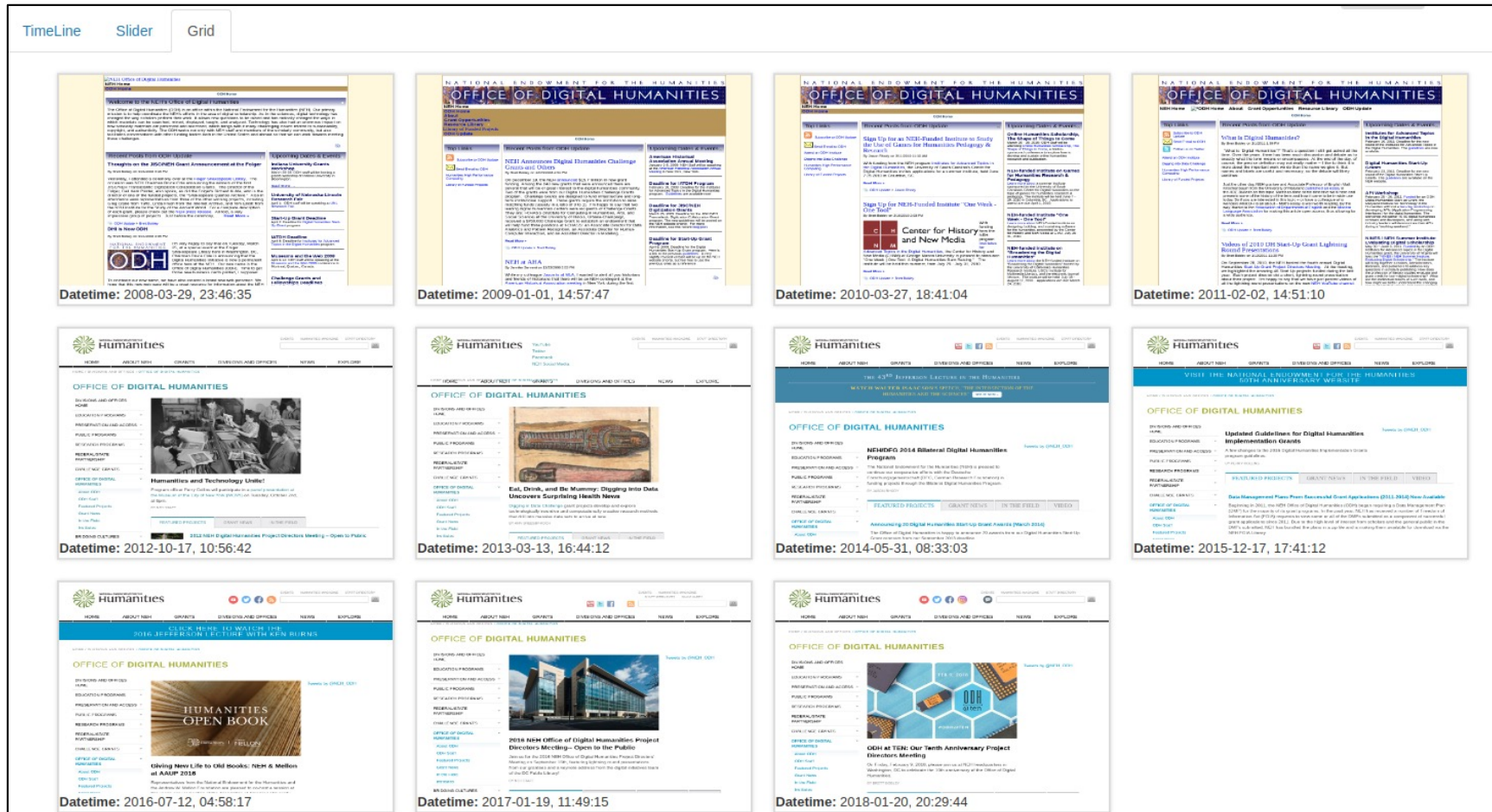
<http://ws-dl.blogspot.com/2017/10/2017-10-16-visualizing-webpage-changes.html>



NATIONAL ENDOWMENT FOR THE  
**Humanities**

"Visualizing Webpage Changes Over Time", 2017-2019, HAA-256368-17

# tmvis – Grid View



<http://ws-dl.blogspot.com/2017/10/2017-10-16-visualizing-webpage-changes.html>



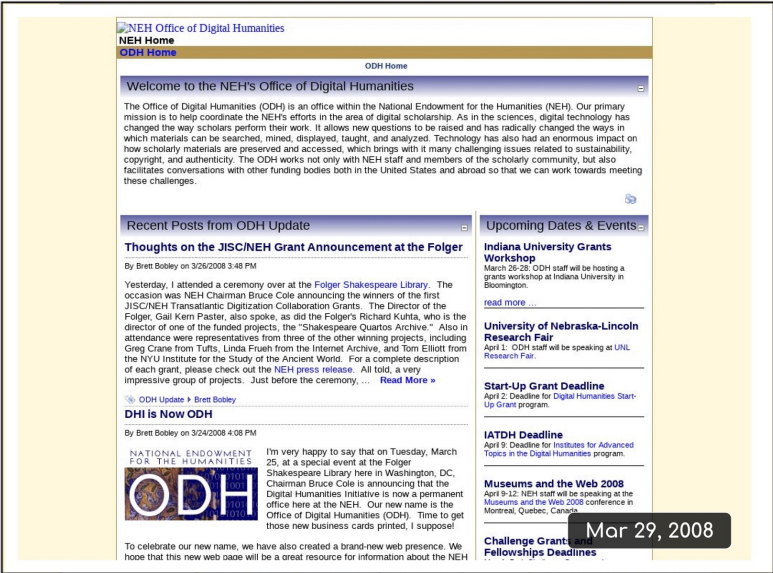
NATIONAL ENDOWMENT FOR THE  
**Humanities**

“Visualizing Webpage Changes Over Time”, 2017-2019, HAA-256368-17

# tmvis- Image Slider View

Swipe over the images to move through time  
Click on an image to view the memento

Play Pause



<http://ws-dl.blogspot.com/2017/10/2017-10-16-visualizing-webpage-changes.html>



NATIONAL ENDOWMENT FOR THE  
**Humanities**

“Visualizing Webpage Changes Over Time”, 2017-2019, HAA-256368-17

# tmvis – Timeline View

Enter URI-R:  
http://www.neh.gov/odh/

Archive-It  Internet Archive

Collection Identifier: 1068 Go

Stats : 502 mementos, 11 Unique Thumbnails

TimeLine Slider Grid

2008 2010 2012 2014 2016 2018

PREV UNIQUE PREV UNIQUE NEXT NEXT

OPEN ARCHIVED PAGE

Datetime: 2008-03-29, 23:46:35

Uses Propublica's TimelineSetter library, <http://propublica.github.io/timeline-setter/>

<http://ws-dl.blogspot.com/2017/10/2017-10-16-visualizing-webpage-changes.html>



NATIONAL ENDOWMENT FOR THE  
**Humanities**

“Visualizing Webpage Changes Over Time”, 2017-2019, HAA-256368-17

# What did we learn from this?

- Webpages can go off-topic through time
- Some mementos aren't captured well
- Some mementos aren't replayed well

# You don't want off-topic mementos in your summary

**PLEASE JOIN OUR TEXT MESSAGE SERVICE!**  
Send INDYOWS (all cases) to (202) 200-4031

## INDYOWS WELCOME

Home | Announcements | Calendar | Forum | Press Releases | Photos | Livestream | Contact Us

### MAJOR ANNOUNCEMENT

**INDYOWS INFORMATIONAL MEET & GREET**  
DATE: Friday, December 8, 2011  
LOCATION: Indianapolis Musicians Local 3 AF of M, 325 N Delaware St, Indianapolis  
TIME: 7 pm - 8 pm  
This will be an informal meeting, with no General Assembly. Come and get to know each other! Light refreshments will be available. Anyone interested in being a part of the Occupy Wall Street Movement may attend.

**INDYOWS GENERAL ASSEMBLY 1**  
DATE: Monday, December 18, 2011  
LOCATION: Indianapolis Musicians Local 3 AF of M, 325 N Delaware St, Indianapolis  
TIME: 7 pm - 8 pm  
This will be the first General Assembly of INDYOWS. Come and be a part of the newest Occupy Wall Street to come to Indiana!

**INDYOWS GENERAL ASSEMBLY 2**  
DATE: Friday, December 23, 2011  
LOCATION: Indianapolis Musicians Local 3 AF of M, 325 N Delaware St, Indianapolis  
TIME: 7 pm - 8 pm  
With the holidays, this General Assembly time and date is subject to change. Please stay tuned to this page for more information!

**INDYOWS GENERAL ASSEMBLY 3**  
DATE: Friday, December 30, 2011  
LOCATION: Indianapolis Musicians Local 3 AF of M, 325 N Delaware St, Indianapolis  
TIME: 7 pm - 8 pm  
Again, with the holidays, this General Assembly time and date is subject to change. Please stay tuned to this page for more information!

Welcome to INDYOWS, the online home of the Indianapolis Occupy Movement!

Datetime: 2012-01-10, 01:41:57

Maintenance Mode

Sorry for the inconvenience.  
Our website is currently undergoing scheduled maintenance.  
Please try back shortly.  
Thank you for your understanding.

Datetime: 2012-04-10, 03:26:34

Ever feel you're in the wrong place  
404 (Page Not Found) Error

Sorry for the inconvenience.  
Our website is currently undergoing scheduled maintenance.  
Please try back shortly.  
Thank you for your understanding.

1. You entered an incorrect URL into your browser's address bar, or
2. You haven't updated content.

If you're a visitor and not sure what happened:

1. If you entered or copied the URL incorrectly or
2. The link you used to get here is faulty.
3. (It's an excellent idea to let the link owner know.)

Datetime: 2012-04-17, 03:26:15

## INDYOWS

The revolution continues worldwide!

Home | Announcements | Calendar | Forum | Press Releases | Photos | Livestream | Contact Us

Welcome Home!

Welcome to INDYOWS, the online home of the Indianapolis Occupy Movement!

We welcome people of all viewpoints, religions, political stances, nationalities, races and sexual orientations and encourage respectful dialogue between those who differ in opinion.

We will be actively working to influence the public on corporate involvement in the government and working towards policy change at the state level. We will work in conjunction with other occupy groups towards concrete change. We welcome all that are ready to take on the monumental and historic task.

### GREAT NEWS!

#### TEXT MESSAGE SERVICE AVAILABLE AGAIN

Back in November when this group started, we had a text messaging service for our members. You could send a message in our special number and you would be subscribed to receive text alerts about INDYOWS. It was run by volunteers and they ended the service "temporarily". That was three months ago!

But now for the GREAT NEWS! Thanks to a group of occupiers we have another text message service available to use! Simply text [@indyows to 23559](#) and you will be subscribed to receive text alerts from INDYOWS.

**INDYOWS OCCUSTOCK 2012 Spring Rally**  
DATE: Saturday May 5, 2012  
ADMISSION: Cash Davidson  
TIME: 11:00AM - 10:30PM  
LOCATION: Johnson County (Indiana) Park Amphitheatre

A day of Fellowship, Peace, Love, Art and Music! Come celebrate the Occupy Movement with INDYOWS! Featuring bands all day long, performance artists, vendors, art, food, and more!

Face painting and so much more! NO protest, NO demonstrations, just a day filled with FUN and FELLOWSHIP!

Bands/Musicians wishing to inquire about performing should contact Nancy Bohannon at [NancyBohannon@indyows.org](#) or Jade Bohannon at [JadeBohannon@indyows.org](#)

Vendors and Artists wishing to rent booth space should contact Karen White at [KarenWhite@indyows.org](#) or Jade Bohannon at [JadeBohannon@indyows.org](#)

EMT/Nurses wishing to donate their time to man the first aid booth should contact Brett Sturgill at [BrettSturgill@indyows.org](#)

Datetime: 2012-04-24, 03:36:58

Forbidden

You don't have permissions to access / on this server.  
Additionally, a 404 Not Found error was encountered while trying to use an ErrorDocument to handle the request.

Apache Server at www.indyows.org Port 80

Datetime: 2012-05-15, 03:47:04

Search Channels

INDYOWS  
This is the official website of the Indianapolis Occupy Movement. We are the only and we are the most active group!

Showing 1 of 1 Channels

Facebook | Twitter | YouTube | More

Datetime: 2012-07-03, 12:18:48

[http://wayback.archive-it.org/2950/\\*/http://www.indyows.org](http://wayback.archive-it.org/2950/*/http://www.indyows.org)

# Identify off-topic mementos with Off-Topic Memento Toolkit (2018)

```
{ "http://wayback.archive-
it.org/1068/timemap/link/http://www.badil.org/" : {
  "http://wayback.archive-
it.org/1068/20130307084848/http://www. badil.org/" : {
    "timemap measures" : {
      "cosine" : {
        "stemmed" : true,
        "tokenized" : true,
        "removed boilerplate" : true,
        "comparison score" : 0.10969941307631487,
        "topic status" : "off-topic"
      },
      "bytecount" : {
        "stemmed" : false,
        "tokenized" : false,
        "removed boilerplate" : false,
        "comparison score" : 0.15971409055425445,
        "topic status" : "on-topic"
      }
    },
    "overall topic status" : "off-topic"
  },
  ...
}
```

Python module

Given a URI-T (TimeMap), identifies off-topic mementos

Option of 8 different similarity measures

OTMT Distribution Page:  
<https://pypi.org/project/otmt/>

OTMT Source Code Page:  
<https://github.com/oduwsdl/off-topic-memento-toolkit>

**Yasmin AlNoamany**, Michele C. Weigle, and Michael L. Nelson, "Detecting Off-Topic Pages Within TimeMaps in Web Archives," *IJDL*, Vol. 17, No. 3, July 2016.

**Shawn Jones**, Michele C. Weigle, and Michael L. Nelson, "The Off-Topic Memento Toolkit," iPres 2018.

 COLUMBIA UNIVERSITY LIBRARIES

"Tools for Managing Seed URIs", 2014-2015, Columbia Libraries Web Archiving Incentive Program



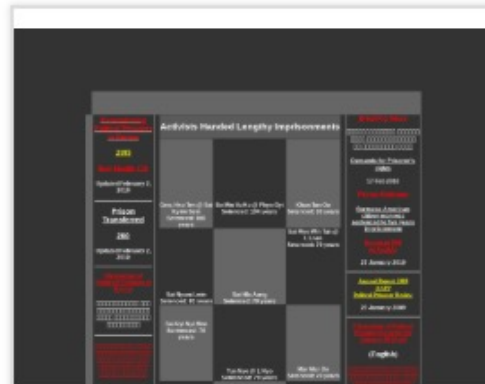
"Combining Social Media Storytelling With Web Archives", 2015-2019, IMLS National Leadership Grant



# You don't want damaged mementos in your summary



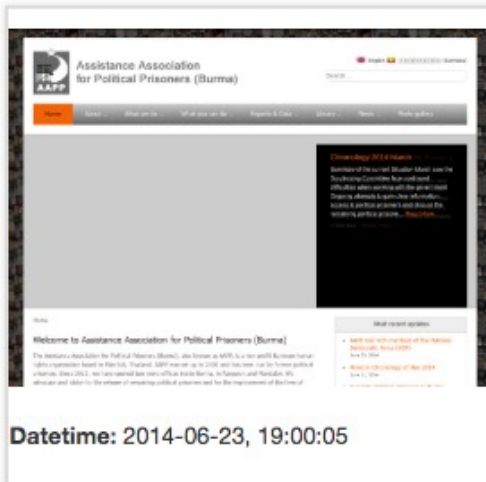
Datetime: 2009-05-14, 15:40:17



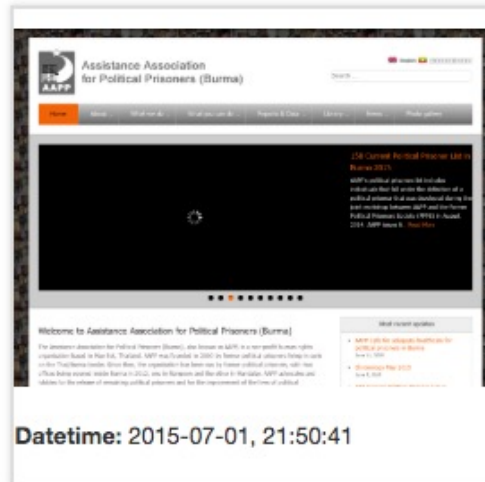
Datetime: 2010-02-23, 15:37:26



Datetime: 2011-06-02, 19:28:38



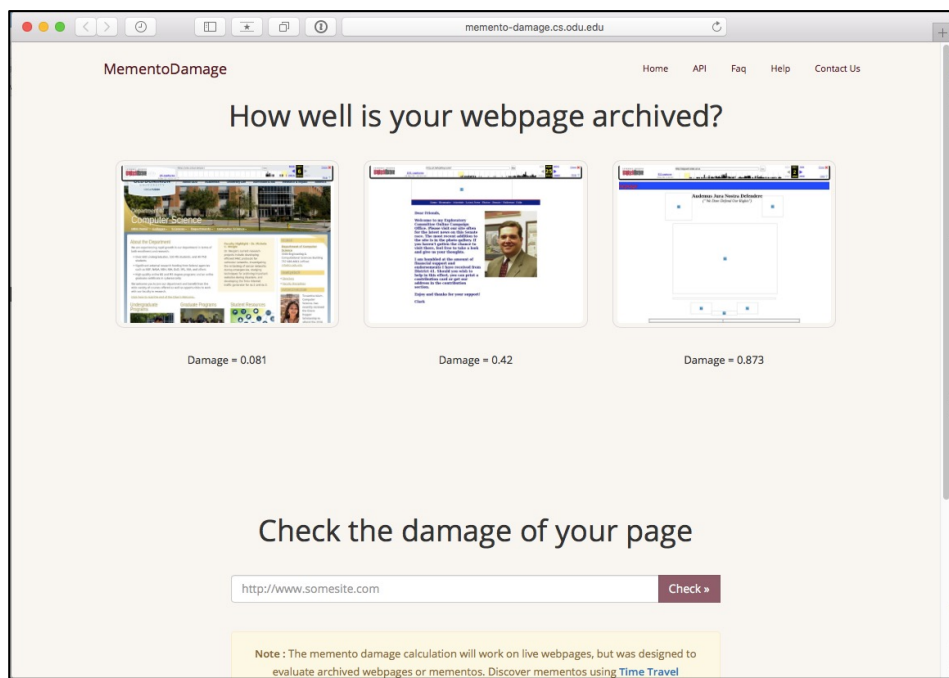
Datetime: 2014-06-23, 19:00:05



Datetime: 2015-07-01, 21:50:41

[https://wayback.archive-it.org/1068/\\*/http://aappb.org/](https://wayback.archive-it.org/1068/*/http://aappb.org/)

# Memento Damage can tell you how damaged your mementos are (2017)



Web service, Docker container

Given URI-M, calculates and analyzes memento damage

Service:

<http://memento-damage.cs.odu.edu>

Github:

<https://github.com/oduwsdl/web-memento-damage>

**Justin Brunelle, Mat Kelly, Hany SalahEldeen, Michele C. Weigle and Michael L. Nelson**, "Not All Mementos Are Created Equal: Measuring the Impact of Missing Resources," *IJDL*, Vol. 16, No. 3-4, September 2015.

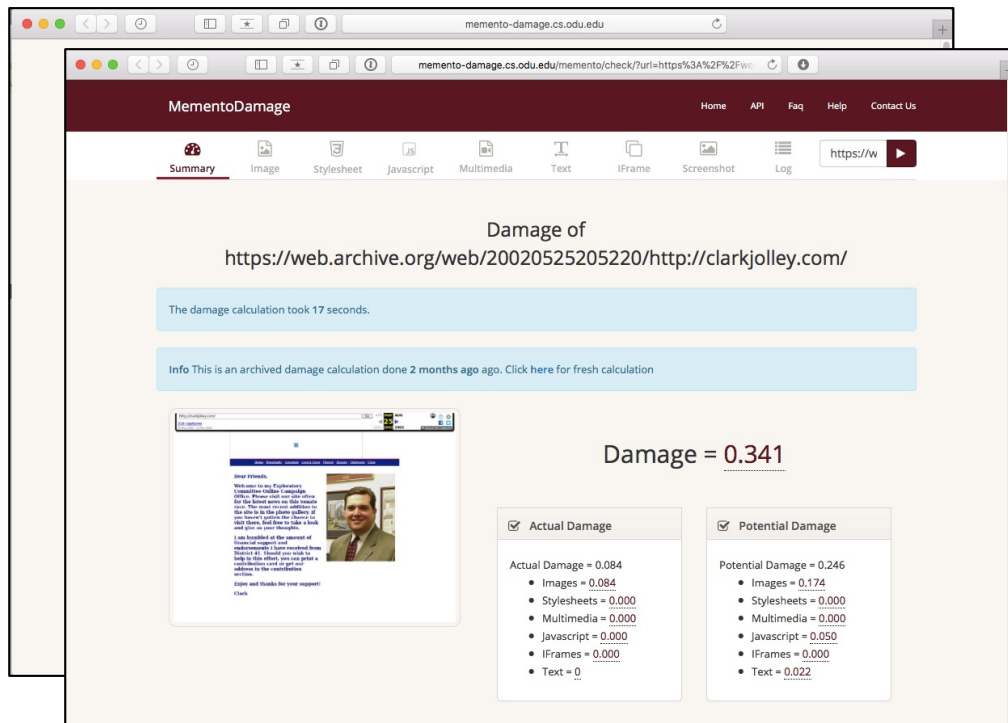
**Erika Siregar**, "Deploying the Memento Damage Service: A Comprehensive Tool for Measuring and Analyzing Damage on Web Archives", ODU MS Project, 2017.

<http://ws-dl.blogspot.com/2017/11/2017-11-22-deploying-memento-damage.html>



"Increasing the Value of Existing Web Archives," 2015-2019, III 1526700

# Memento Damage can tell you how damaged your mementos are (2017)



Web service, Docker container

Given URI-M, calculates and analyzes memento damage

Service:

<http://memento-damage.cs.odu.edu>

Github:

<https://github.com/oduwsdl/web-memento-damage>

**Justin Brunelle, Mat Kelly, Hany SalahEldeen, Michele C. Weigle and Michael L. Nelson**, "Not All Mementos Are Created Equal: Measuring the Impact of Missing Resources," *IJDL*, Vol. 16, No. 3-4, September 2015.

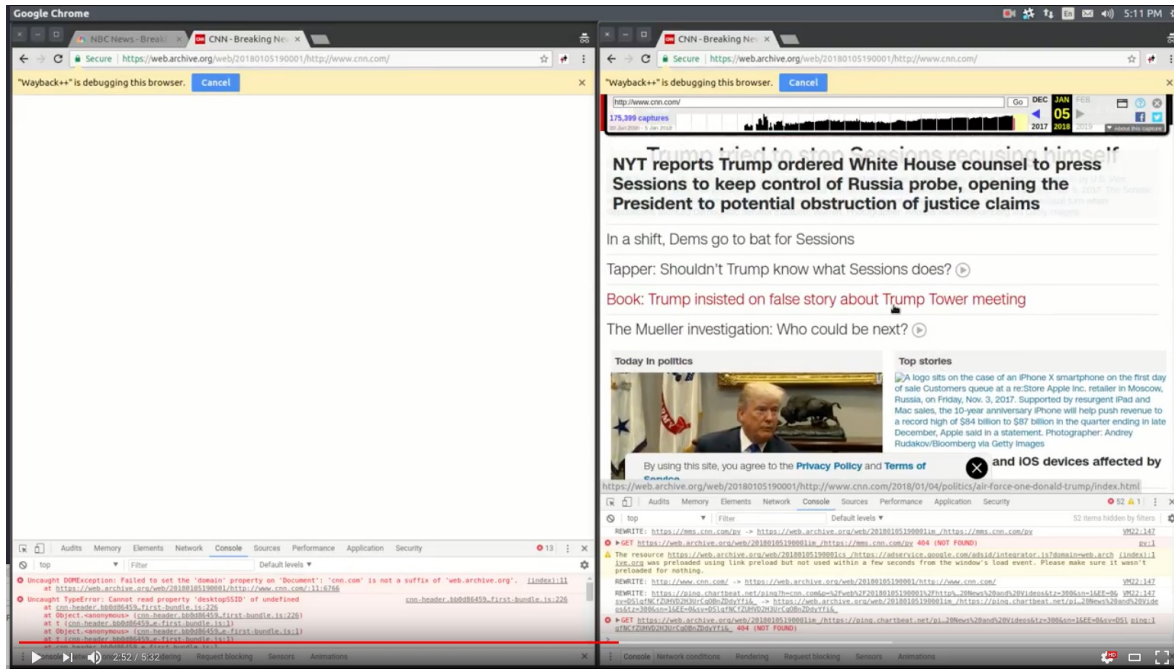
**Erika Siregar**, "Deploying the Memento Damage Service: A Comprehensive Tool for Measuring and Analyzing Damage on Web Archives", ODU MS Project, 2017.

<http://ws-dl.blogspot.com/2017/11/2017-11-22-deploying-memento-damage.html>



"Increasing the Value of Existing Web Archives," 2015-2019, III 1526700

# Wayback++ uses client-side rewriting to fix replay-based damaged mementos (2018)



Chrome, Firefox extensions

<https://github.com/N0taN3rd/WaybackPlusPlus>

<https://www.youtube.com/watch?v=lidyidcaVXHw>

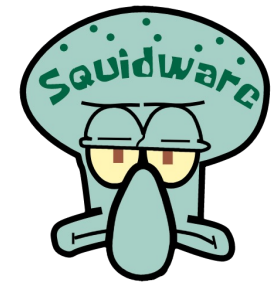
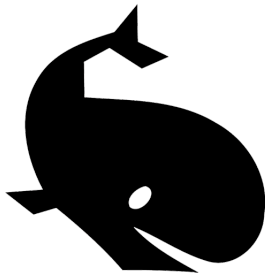
<http://ws-dl.blogspot.com/2017/01/2017-01-20-cnncom-has-been-unarchivable.html>

**John Berlin, Michael L. Nelson, and Michele C. Weigle, "Swimming In A Sea Of JavaScript, Or: How I Learned To Stop Worrying And Love High-Fidelity Replay," WADL 2018.**

<http://ws-dl.blogspot.com/2018/04/2018-05-01-high-fidelity-ms-thesis-to.html>

# Where does this take us?

# We've developed a lot of tools



# But, can a full professor use them?



Fred Brooks says:

Two of our criteria for success in a tool are:

- It must be so *easy* to use that a full professor *can* use it, and
- It must be so *productive* that full professors *will* use it.

Frederick P. Brooks, Jr. 1996. The computer scientist as toolsmith II. Commun. ACM 39, 3 (March 1996), 61-68.

# So, let's think bigger

- In a world where the web browser *is* the Internet, how can we make web archives ubiquitous?



# So, let's think bigger

- In a world where the web browser *is* the Internet, how can we make web archives ubiquitous?
- Bring web archives to the browser - *natively*

Michele C. Weigle, Michael L. Nelson, Martin Klein, and Herbert Van de Sompel, “The Case for Memento-Aware Browsers”, 2017



# What if browsers could natively identify mementos?

- Look for Memento-Datetime header in HTTP response

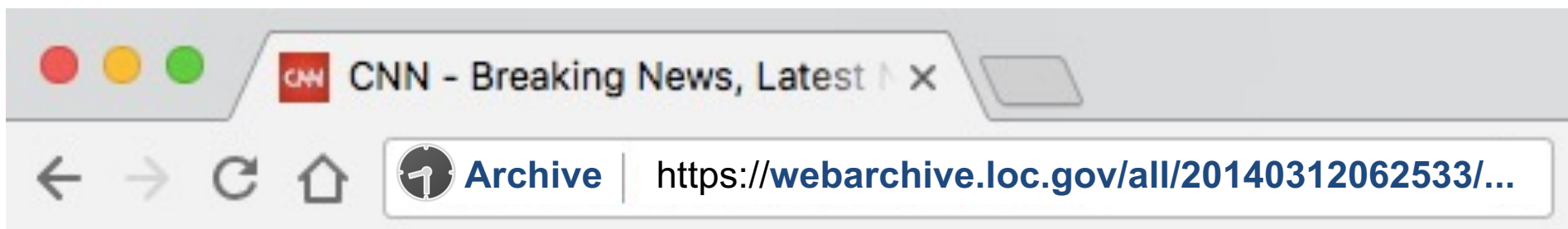
`Memento-Datetime: Tue, 08 May 2012 11:24:30 GMT`

- Use client-side rewriting (Emu) to improve replay
- Use native UI elements to annotate composite mementos

# Identify mementos in the address bar



# Identify mementos in the address bar



Could also identify non-HTML mementos (images, PDF, etc.)

# Identify temporal inconsistencies

The screenshot shows a web browser window with the address bar displaying "http://web.archive.org/web/20050601025530/..". The page content includes the CNN logo, the New York Times logo, and various advertisements. A red box highlights a Dell advertisement for a desktop and notebook. The advertisement text reads: "\$373 \$299 Desktop & \$699 \$560 Notebook from Dell Home. Offer ends 6/1." Below the text is an image of a Dell desktop computer system (monitor, keyboard, mouse, and tower) and the Dell logo. A button labeled "CLICK FOR DETAILS" is visible at the bottom right of the advertisement.

Scott Ainsworth, <http://ws-dl.blogspot.com/2015/12/2015-12-08-evaluating-temporal.html>

# Identify temporal inconsistencies

The screenshot shows a browser window with the address bar containing "http://web.archive.org/web/20050601025530/..". The page header includes the New York Times logo and a date stamp: "UPDATED TUESDAY, MAY 31, 2005 10:55 PM ET". A red box highlights a Dell advertisement for a desktop and notebook, with a mouse cursor pointing to it. A red callout box above the ad says "+ 5 Years, 11 months (Apr 6, 2011)". To the right, a "Most E-Mailed" list contains five items, with the fifth item, "C.I.A. Expanding Terror Battle Under Guise of Charter Flights", highlighted in blue. The word "Business" is visible at the bottom right of the page content.

Scott Ainsworth, <http://ws-dl.blogspot.com/2015/12/2015-12-08-evaluating-temporal.html>

# What if browsers could natively interact with Memento aggregators?

- Alert users of unarchived pages as they browse
- Provide UI elements to summarize and access past versions of the current webpage
- Integrate web archives and the past web into “New Tab View”

# What if browsers could natively interpret and replay WARC?

- Users could share WARCs
- Recipient could open the WARC directly in their browser
- WARC.js (ala PDF.js for WARCs)



# What if browsers could natively create mementos?

- Push to public web archives
- Create local WARC

Just as easily as taking a screenshot  
*or maybe along with taking a screenshot*

**Conspirador Norteño** @conspirator0 · May 26  
This account changed names sometime during that four-year gap. Based on this snapshot from the Internet Archive, @veteranhnak was @louiebaur in 2012. Curiously, there's a @louiebaur account now with tweets from 2011, indicating that at least two accounts were renamed.

Google TV's second wind: New devices coming from LG, Samsung, Vizio and Sony dvr:/3wFF

**Veteran Hawk** @veteranhawk · 5 Jan 2012  
Mobile ac network/Millennial Media files for (IPO dvr:/3wFF)

**Veteran Hawk** @veteranhawk · 5 Jan 2012  
Replying to @Gitteree  
@Gitteree I sure am did it work?

**Veteran Hawk** @veteranhawk · 5 Jan 2012  
Replying to @nancyanteb  
@nancyanteb LOL totally

**Veteran Hawk** @veteranhawk · 5 Jan 2012  
Replying to @BobWarren  
@BobWarren You to wait!

**Veteran Hawk** @veteranhawk · 5 Jan 2012  
I haven't heard Casey Anthony in awhile

**louiebaur** Louie Baur  
Google TV's second wind: New devices coming from Samsung, Vizio and Sony dvr:/3wFF

**louiebaur** Louie Baur  
Mobile ac network/Millennial Media files for (IPO dvr:/3wFF)

**louiebaur** Louie Baur  
@Gitteree I sure am did it work?

**louiebaur** Louie Baur  
@nancyanteb LOL totally

**louiebaur** Louie Baur  
@BobWarren You to wait!

**Bob Warren** Bob Warren  
The Only Word Linkedin Game: @nyan.com/2b6tz

**louiebaur** Louie Baur  
I haven't heard Casey Anthony in awhile

2 9 40

**Conspirador Norteño** @conspirator0 **Follow**

(Archive link the screenshot in the previous tweet was taken from)  
[web.archive.org/web/2012010617...](https://web.archive.org/web/2012010617...)

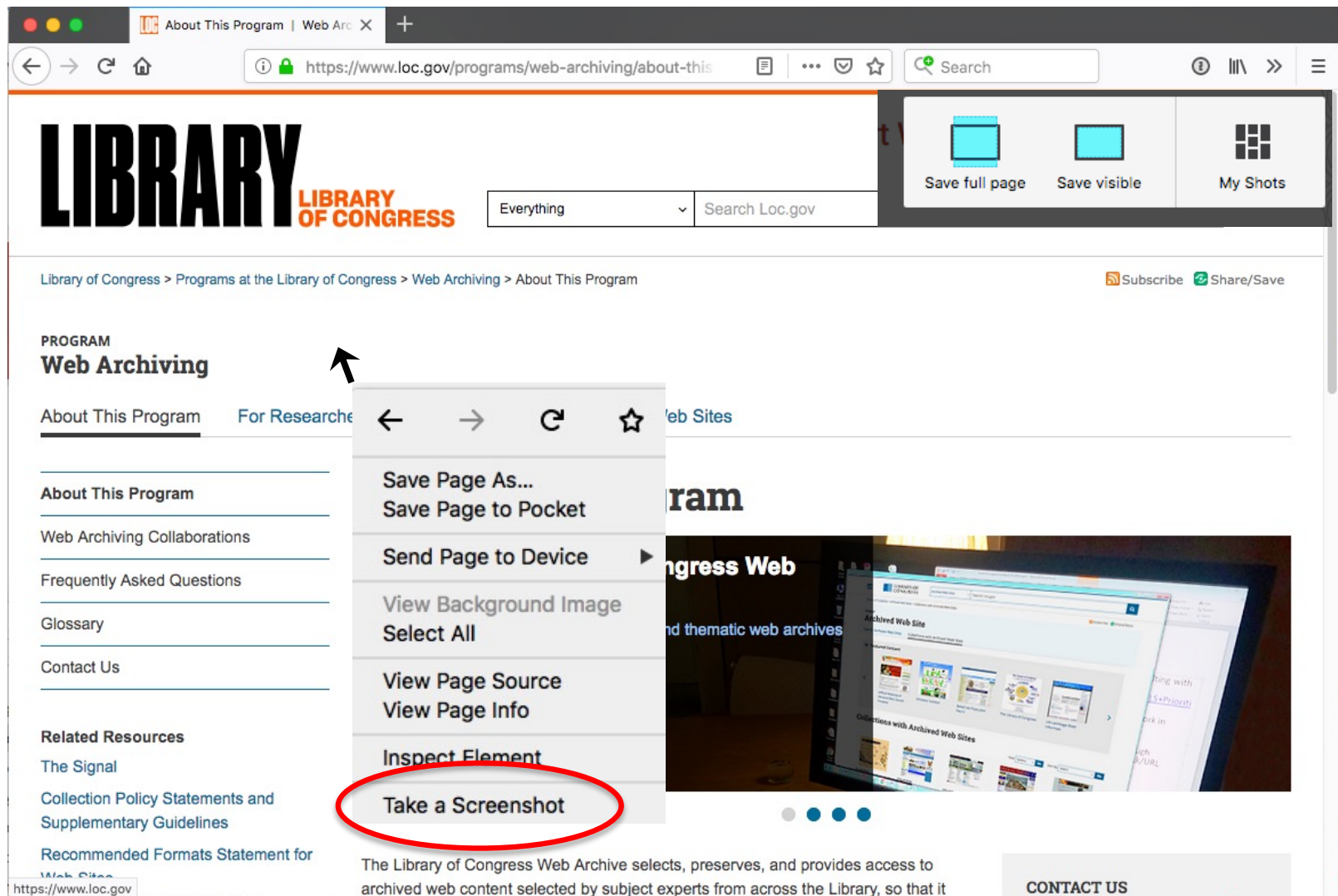
4:33 PM - 26 May 2018

4 Retweets 27 Likes

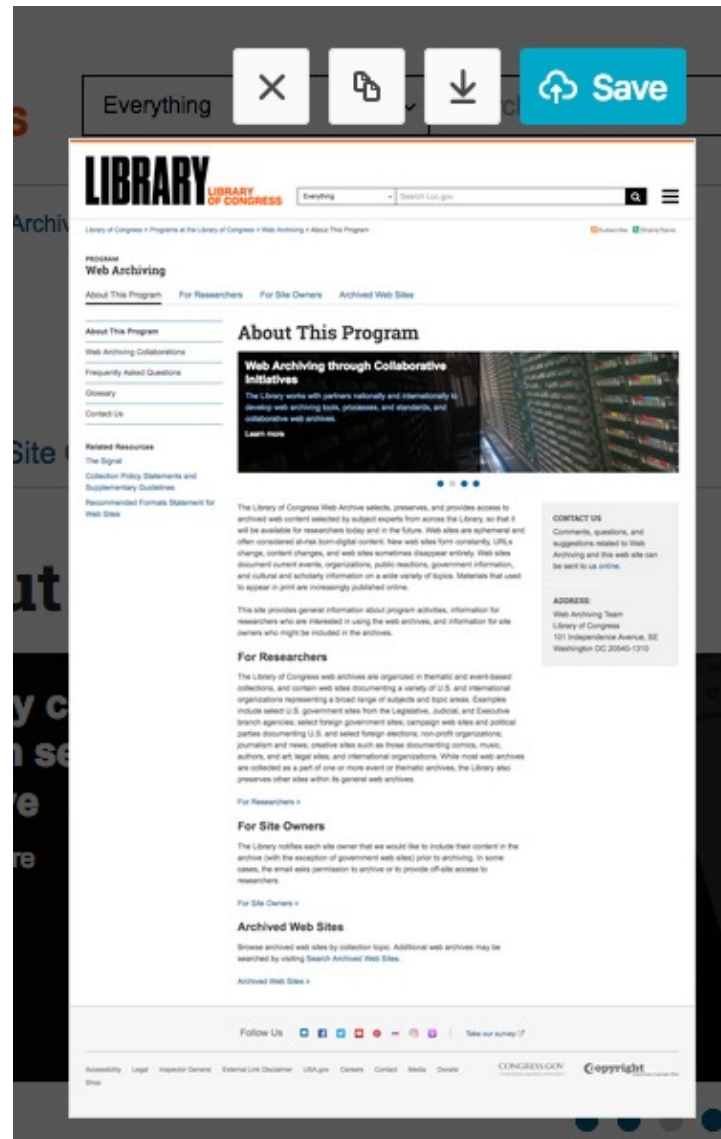
1 4 27

<https://twitter.com/conspirator0/status/1000475042017366017>

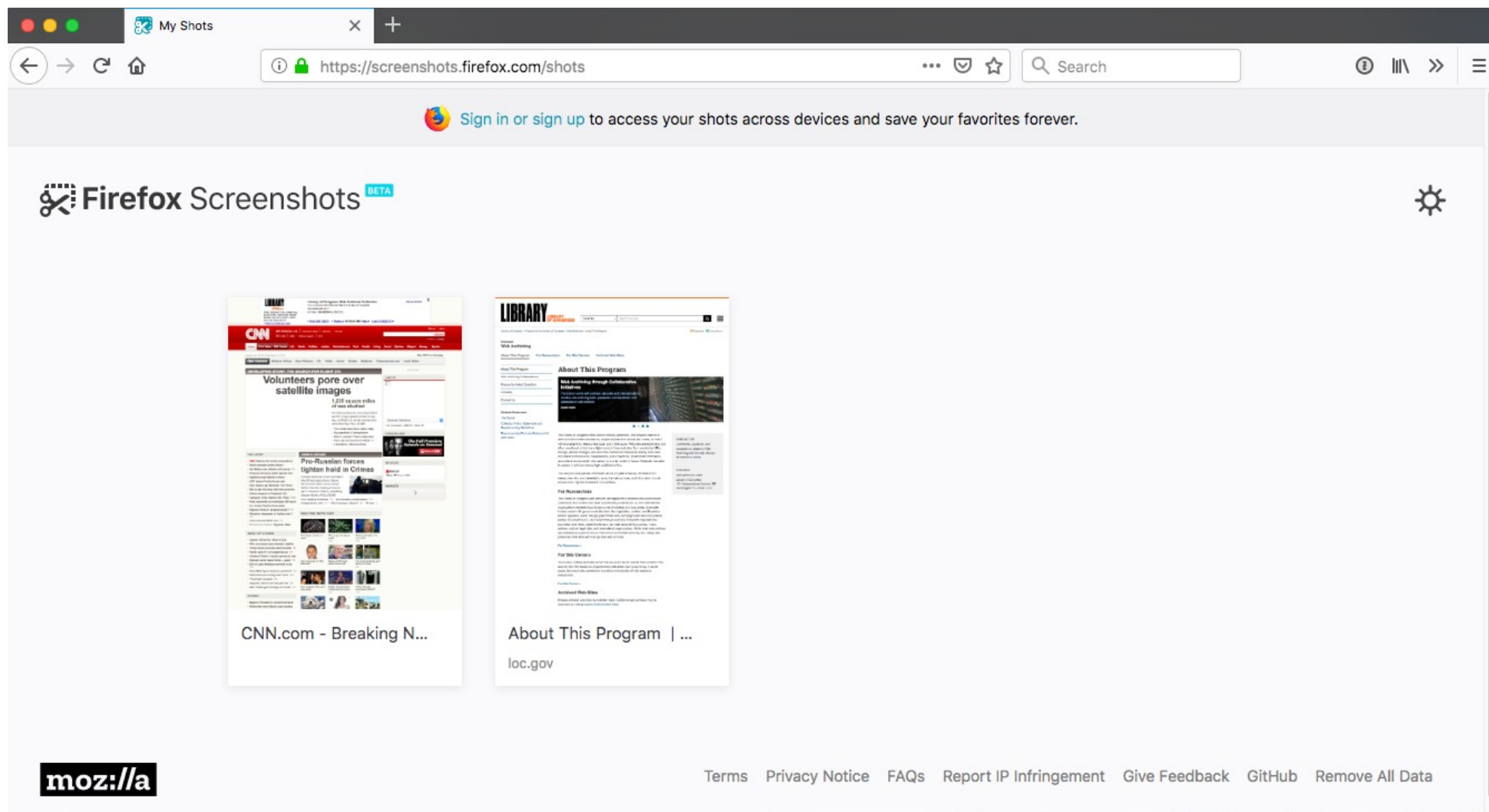
# Firefox Quantum has brought screenshots natively to the browser



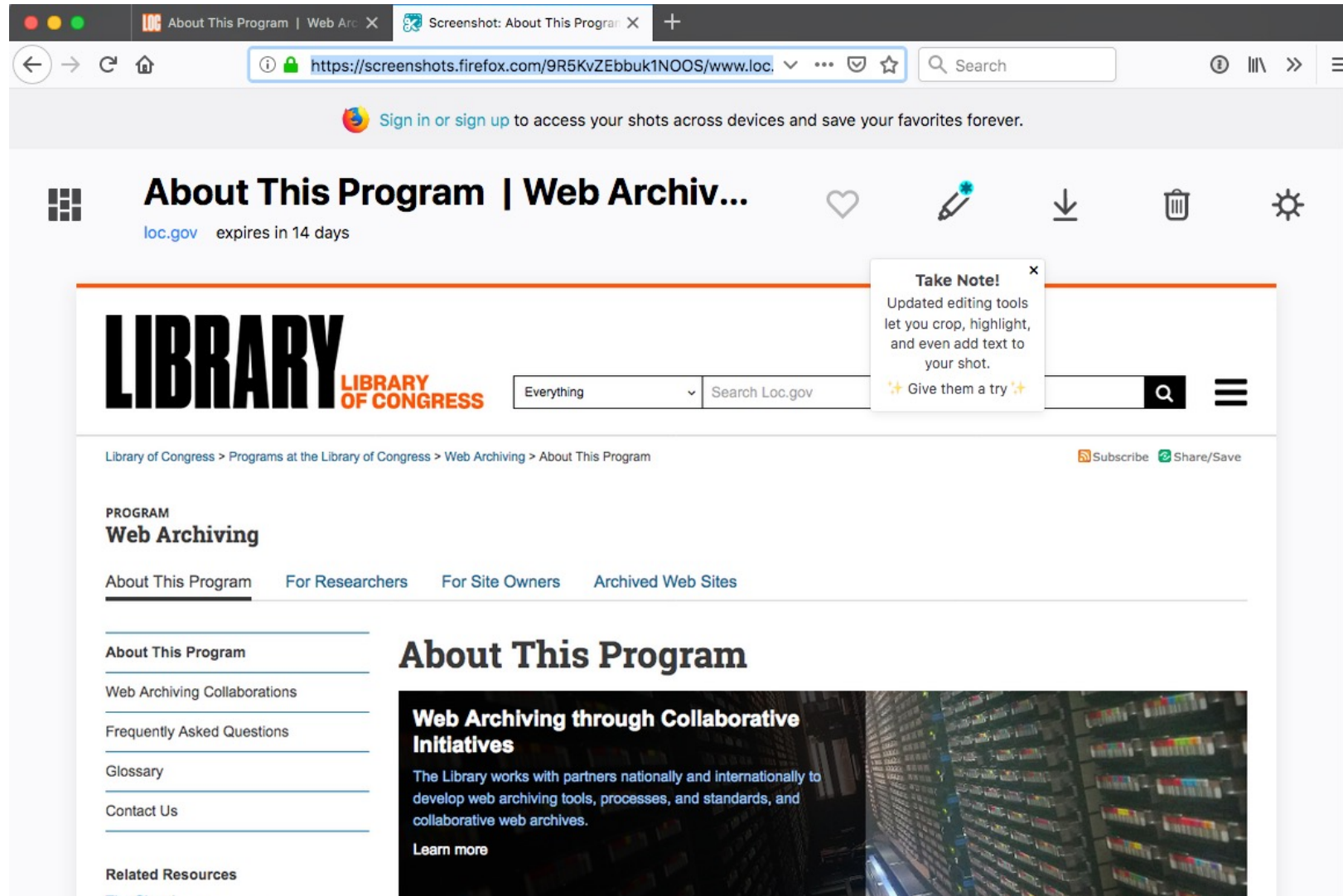
# Saving full page screenshot



# Screenshots can be saved in the Mozilla cloud



# Screenshots have a URI



<https://screenshots.firefox.com/9R5KvZEbbuk1NOOS/www.loc.gov>

# What if these screenshots were Memento-enabled?

- Provide Memento HTTP headers for the screenshots
- Implement Memento datetime negotiation for the entire screenshot cloud service

# We could build a crowd-sourced archive of screenshots

- Take screenshot and save to Memento-enabled screenshot cloud
- Option to push live webpage to archive at same time
- Then we have both an archived page and a screenshot of the page from very close to the same datetime

# What about bookmarks?

The screenshot shows a web browser window with a tab titled "Home | Michele C. Weigle". The address bar shows the URL "www.cs.odu.edu/~mweigle/". The main content area displays the profile of Michele C. Weigle, an Associate Professor at Old Dominion University (ODU). The profile includes a photo, contact information, and sections for "Teaching/Admin News" and "Research News".

Overlaid on the right side of the browser is a "Bookmark added" dialog box. The dialog box contains the following information:

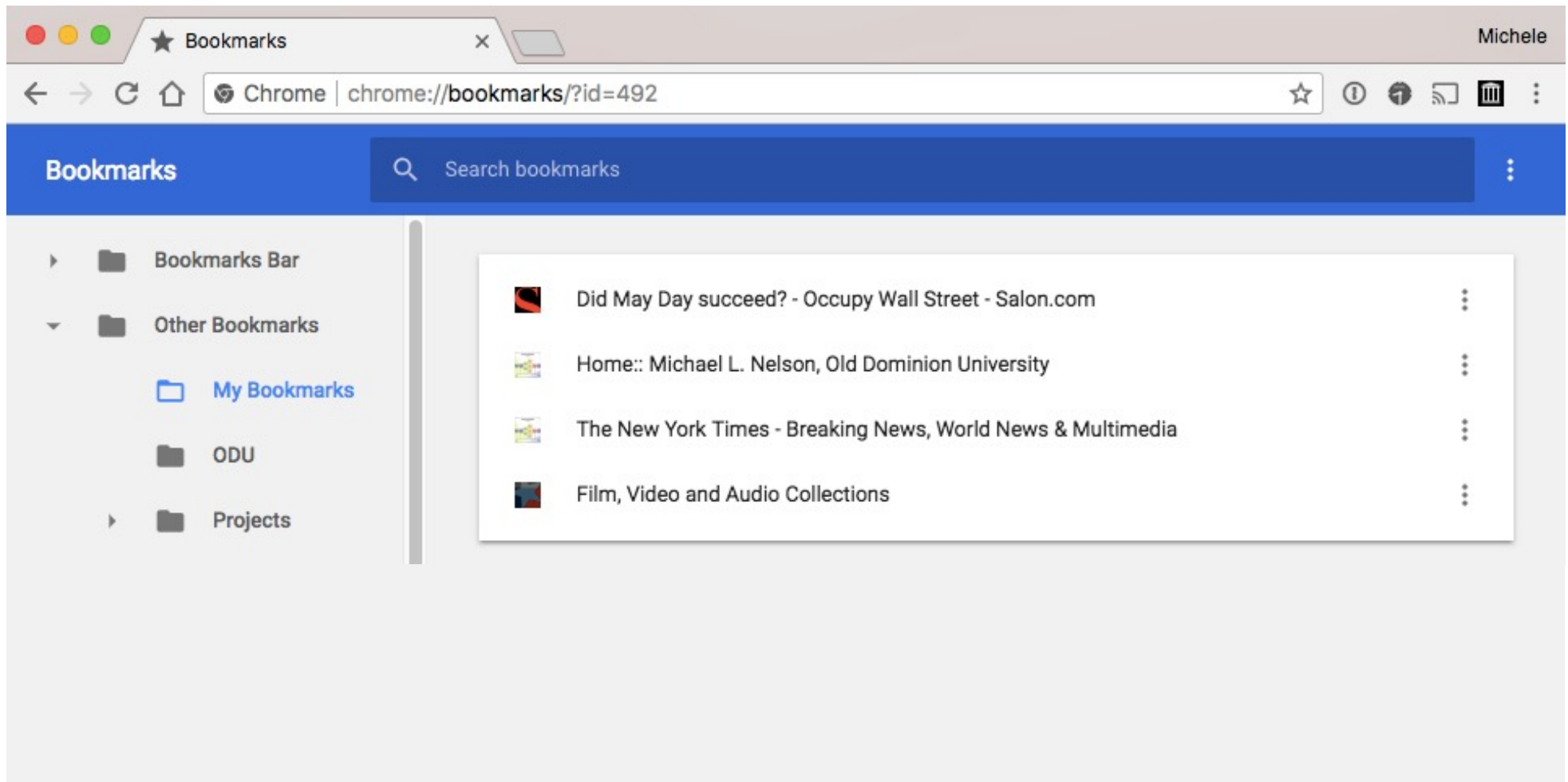
- Name:** Home | Michele C. Weigle
- Folder:** My Bookmarks
- Submit to public web archives:** A toggle switch is turned on.
- Buttons:** More..., Remove, and Done.

Below the dialog box, a message reads: "local archive saved to ~/Library/WebArchive/".

Bookmarking becomes archiving



# Viewing a bookmark becomes an opportunity to interact with archives



# Memento Embeds for bookmark view

The screenshot shows a Chrome browser window with the address bar at `chrome://bookmarks/?id=492`. The page title is "Bookmarks" and the user's name is "Michele". The left sidebar contains a "Bookmarks" menu with options: "Bookmarks Bar", "Other Bookmarks", "My Bookmarks", "ODU", and "Projects". The main content area displays four memento embeds:

- Did May Day succeed? - Salon.com**  
Member of [Archive-It Collection Occupy Movement 2011/2012](#)  
Live web resource is Unavailable  
Yesterday's Occupy reboot mobilized a diverse group of people, but reverted to familiar tactics in the end  
[WWW.SALON.COM @ 2012-05-08T03:32:01Z](#)
- Home:: Michael L. Nelson, Old Dominion University**  
Live web resource is Available  
About Me... I joined the Computer Science department at Old Dominion University in 2002. I worked at NASA Langley Research Center from 1991-2002. Through a NASA fellowship, I spent the 2000-2001 academic year at the School of Information and Library Science, University of North Carolina at Chapel...
- The New York Times - Breaking News, World News & Multimedia**  
Live web resource is Available  
Find breaking news, multimedia, reviews & opinion on Washington, business, sports, movies, travel, books, jobs, education, real estate, cars & more.  
[NYTIMES.COM:80 @ 2005-06-01T02:55:30Z](#)
- Film, Video and Audio Collections**  
Member of [Archive-It Collection IT Historical Resource Sites](#)  
Live web resource is Unavailable  
Industry On Parade Film Collection, 1950-1960  
[AMERICANHISTORY.SI.EDU @ 2010-05-20T21:22:53Z](#)

# Open live web, local memento, or public memento

The screenshot shows a Chrome browser window with the address bar displaying `chrome://bookmarks/?id=492`. The page title is "Bookmarks" and the user's name "Michele" is visible in the top right. The bookmarks are organized into folders: "Bookmarks Bar", "Other Bookmarks", "My Bookmarks", "ODU", and "Projects".

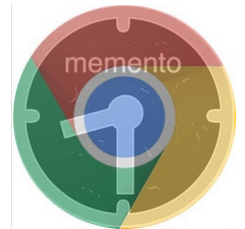
Four bookmark cards are visible:

- Did May Day succeed? - Salon.com**: Member of Archive-It Collection Occupy Movement 2011/2012. Live web resource is Unavailable. Description: Yesterday's Occupy reboot mobilized a diverse group of people, but reverted to familiar tactics in the end. URL: `WWW.SALON.COM @ 2012-05-08T03:32:01Z`.
- Home:: Michael L. Nelson, Old Dominion University**: Live web resource is Available. Description: About Me... I joined the Computer Science department at Old Dominion University in 2002. I worked at NASA Langley Research Center from 1991-2002. Through a NASA fellowship, I spent the 2000-2001 academic year at the School of Information and Library Science, University of North Carolina at Chapel... URL: `WWW.CS.ODU.EDU @ 2018-05-15T13:00:56Z`.
- The New York Times - Breaking News, World News & Multimedia**: Live web resource is Available. Description: Find breaking news, multimedia, business, sports, movies, travel, books, jobs, education, ... URL: `NYTIMES.COM:80 @ 2005-06-01T02:55:30Z`.
- Film, Video and Audio Collections**: Member of Archive-It Collection IT Historical Resource Sites. Live web resource is Unavailable. Description: Industry On Parade Film Collection, 1950-1960. URL: `AMERICANHISTORY.SI.EDU @ 2010-05-20T21:22:53Z`.

A context menu is open over the "The New York Times" bookmark, showing the following options:

- Edit
- Copy URL
- Delete
- Open on live web
- Open local memento
- Open public memento

# It's time for browsers to be Memento-aware



- Web archives have gone mainstream.
- We've learned a lot by building tools to enable personal use of web archives.
- These ideas need to be integrated directly into browsers for general public use.