

**Special interest session: Web archiving**  
Pecha Kucha and discussion of emerging topics in web archiving  
NDIIPP/NDSA Partners Meeting Workshop  
Wednesday, July 20, 2011  
11:15 a.m.

**Presenters:** Abbie Grotke (Library of Congress); Kris Carpenter Negulescu (Internet Archive); Mat Kelly, Vivens Ndatinya, and Michael Nelson (all Old Dominion University); Victoria Priester (Library of Congress)

**Attendees:** 25

**Highlights of the topics covered during the meeting.**

Abbie Grotke gave an overview of the Library's event harvesting and the benefits of cooperative archiving, which includes being able to react quickly and archive more content. Some of the collaborative archive projects mentioned included September 11, Hurricane Katrina, and the earthquakes in Haiti and Japan.

Vivens Ndatinya and Mat Kelly talked about some new developments in being able to capture social media. Vivens explained that through a 'resolver service', users don't have to worry about YouTube music videos that have been removed. This service can find the metadata for the video and thus find additional copies that might be on the web. Mat assisted in developing a Firefox add-on that will help users archive their Facebook page. Currently, not all of a Facebook page is archived but with this program the user will be able to archive their friends list and group pages.

Michael Nelson and Victoria Preister gave more analytical presentations. Michael explained some testing that had been done to see just how much of the web was actually being archived. The bottom line was that much of the web is archived, but it depended on what part of the web was being discussed, in particular search engines. Victoria went through an election analysis that had begun in 2002. By following a particular candidate that has run for office every couple of years, users can see how the social media portion has evolved, how e-donations play a large part, and how many candidates actually paid for a website in 2008 as opposed to 2002.

Kris Carpenter Negulescu discussed more of the technical tools that are helping to do better web archiving. For example, a program called *PIG-latin* is being used now for better scripting. With Web Archive Transformation (WAT), users will be able to extract everything that they would normally get from the ARC/WARC file.

**Slide presentations.**

- Workshop3\_Slides-LC.pdf
- Workshop3\_Slides-ODU-A.pdf
- Workshop3\_Slides-ODU-B.pdf
- Workshop3\_Slides-ODU-C.pdf
- Workshop3\_Slides-IA.pdf