Archiving Social Media in Senators' Offices

Records created as a result of work conducted for the Senator (excluding committee records) are the Senator's personal property and should be retained as part of his or her archival collection (H. Con. Res. 307, 110th Congress). A record is not defined by its format; thus communications and information distributed on social media sites as well as the responses and interactions generated by such postings are all examples of records that should be appraised for inclusion in the archival collection. A body of tweets retained over time can illuminate debates, insights into the progression of issues, and consistency or changes in position on a given topic. Public responses and social media communications can provide valuable contextual understanding of the Senator's record. The rapid rise in the use of social media makes it even more important to adopt a practical approach to managing and archiving such records. It is important to remember that social media is dynamic with content that is ever-changing and information that may be picked up by others for use and reuse. Often it is hosted on third party sites that could limit future access to materials posted. All of this adds to the imperative of ensuring that an accurate and authentic 'original' copy of information is preserved and archived as a record.

The records management policy of the office should address management of social media records. More often than not there are clear parallels between social media records and the existing office records policy. The distinction is the interactive and responsive nature of social media and consideration of the value of these interactions needs to be made by those responsible for such records. This responsibility often belongs to communications and press staff. It is interesting to note that nine Senators from the 112th Congress have staff with job titles using the word 'media', most often 'New Media Director'. Whatever their title, office staff using social media tools for official purposes need to work with their office's designated records manager to determine which records should be archived in accordance with office records management policy.

NARA guidance to federal agencies on managing records in Web 2.0/Social Media Platforms (NARA Bulletin 2011-02) identifies some of the challenges associated with such records and all of these points merit careful consideration:

- Public expectations that all web content is both permanently valuable and accessible
- Content located in multiple places
- Recordkeeping in a collaborative environment
- Ownership and control of data that resides with a third party
- Interactive content management
- Identification of record series
- Implementation of records disposition schedules, including the ability to transfer and permanently delete records or perform other records management functions
- Capture of frequently updated records

For material generated by your office keep in mind that traditional principles of good recordkeeping remain important. Do not upload anything to the Web without archiving a local copy first. Saving an archival quality copy of all photos and videos posted online is particularly important. It means never having to resort to retrieving your own media by downloading it from a networking site. Such downloads can be a potential source of viruses and malware and retrieved content may not be available in the original high resolution format used at the time of upload. For other types of records, the accuracy of an original post can be verified by keeping an archival copy of everything posted. Creating archival copies assures that the office maintains control of the longevity and availability of these materials and simply makes sense.

Keeping archival copies of records posted online is only part of the equation. Equal consideration should be given to setting up a schedule to archive these sites on a regular basis to capture the responses and interactions generated by material posted. Offices are also likely to be tracking references to the Senator across various social media sites for current analytics that may become part of the Senator's permanent record. It is unwise to rely on any large commercial organization or third party hosting site for ongoing access to valuable historical materials. Should third party sites decide to restrict access to posted materials, any photos, videos, online interactions and references might no longer be available to the office. Once offices determine what social media records will be collected as part of their records management policy, they are advised to create a record schedule for that collection, to retain saved records on their own servers and to donate this along with the member's collection to an archival repository.

There are a myriad of tools available to back up, capture, and download different forms of social media. These may take the form of freeware, open source applications, add-ons or fee-based options. Some provide backup in the cloud while others can be installed on local servers. Many of these tools have emerged in response to e-discovery and financial regulation requirements with a recent trend towards using one approach to back up all forms of social media in use. There are several services available that provide one-stop solutions; some offer a free solution in addition to a fee-based "pro" version while others offer a free trial to test the usefulness of the service. Many of these include website archiving in conjunction with social media archiving. Some sites offer tools that are primarily for analysis, but contain back-up solutions that may be of interest. At this point Senate offices are most likely to employ Facebook, Twitter, YouTube and Flickr and there are ways to handle each of these individually as well.

Facebook: Best practice recommendations for Facebook pages include posting contact information, naming one individual to maintain the page on a daily basis, and capturing the interactive experience of Facebook. For personal pages Facebook offers a feature titled "Download Your Information" available under Account Settings and this enables download of a zip file of all profile information, such as photos, wall posts, messages, videos, and more. The zip file can be opened and viewed as an HTML page with simple links to all of these things. At this time, Facebook does not offer this feature for fan pages, but several resources listed in the Appendix below, including Backupify and SocialSafe, do include fan pages in their social media

backup. Always be sure to keep a locally archived copy of photos and videos uploaded to Facebook and label them with dates and identification of subjects and events.

Twitter: The Library of Congress announced in April of 2010 that Twitter had donated their entire archive of public tweets dating back to March of 2006 to the Library, but offices should not rely on this as an archival back-up. An FAQ posted on the Library of Congress blog states "Private account information and deleted tweets will not be part of the archive. Linked information such as pictures and websites is not part of the archive, and the Library has no plans to collect the linked sites." Twitter recently announced that they are building a tool that will allow users to download their own tweets, but it is uncertain when that will be available. Although members can access their own tweets, Twitter currently limits users to obtaining their last 3200 tweets and the ability to search tweets historically often goes back less than two weeks. Best practice recommendations include assigning one person to maintain the account daily and to track posted and received tweets on at least a weekly basis. For Senators who tweet, regular archiving of these materials may contribute an added dimension to their historical record. Tweets can be saved by date range and if the tweets are on one topic or thread, a subject heading can be added to the saved file. Check out Politwoops at: http://politwoops.sunlightfoundation.com/ for some interesting tweet viewing, Hosted by the Sunlight Foundation, this is a collection of deleted tweets from U.S. politicians.

YouTube: There are applications that enable download of video from YouTube, including a Firefox Add-on, but the downloads may be in a format that will require conversion or download of an FLV viewer in order to view them. Always be aware that downloaded videos may introduce viruses or malware to your computer. For video content that originates in the office it is best to retain an archival copy prior to upload. File videos in their own directory and label them with subject, date, and name of event. If you decide to capture comments posted in response to a video as screen shots, plan to save these files in the same directory as the videos. Label files with subject, date, and name of event. Available applications can be used to capture video content posted by others, but it is important to become familiar with YouTube's terms of service before first.

Flickr: Flickr enables one to upload and make publicly or privately available many photos at a time. These might be from a campaign event or a speech and the ability to apply multiple tags to support easy access and flexible organization of groups of photos within a larger collection add to Flickr's attraction as a place to store photos. But keep in mind that Flickr is not an archival storage site and may not be with us ten or twenty years from now. It has been recently observed that Facebook and Pinterest are being used with greater frequency to host uploaded photos. Flickr will delete and limit the ability to tag or edit older photos in free accounts. Best practice is to retain an archival copy locally before uploading to Flickr. Label photos with subject, date, and name of event.

Appendix

The following table includes various tools and applications, both fee-based and free that are used to back up or archive types of social media and web pages. This is not a comprehensive list nor is it an endorsement of any item. It is for informational purposes to show the types of applications available.

Title	URL	Social Media Platform	Notes
BackUpmyTweets	http://backupmytweets.com/	Twitter	Free and Premium edition
TweetBackup	http://tweetbackup.com/	Twitter	Daily Backup - Powered by backupify
Tweet Nest	http://pongsocket.com/tweetnest/	Twitter	Installed on your web server
TWEETSCAN Backup	http://www.tweetscan.com/data.php	Twitter	Has an historical index going back to December of 2007
Twistory	http://twistory.net/	Twitter	Adds your twitter backlog to your calendar - free or pro versions
Grabstream	https://addons.mozilla.org/en-US/firefox/addon/grabstream/?src=api	You Tube	Firefox add-on
youtube-dl	http://rg3.github.com/youtube-dl/	You Tube	This is a command line program that uses the Python interpreter
You Tube Downloader	https://addons.mozilla.org/en-US/firefox/addon/youtube-downloader/?src=search	You Tube	Firefox add-on
Easy You Tube Video Downloader	https://addons.mozilla.org/en-US/firefox/search/?q=you+tube&appver=13.0.1&platform=windows	You Tube	Firefox add-on
You Tube File Hack 2	http://blog.unlockforus.org/2010/02/how-to-archive-youtube-videos-in.html	You Tube	Works w/ Windows 7
Aviosoft YouTube Downloader 2.0.0.1	http://www.softpedia.com/get/Internet/Download-Managers/Aviosoft-YouTube-Downloader.shtml	You Tube	Paste video links into software to download; can convert to different formats
FlickrEdit	http://sunkencity.org/flickredit	Flickr	A Java desktop application
FlickrDownloadr	http://flickrdownloadr.codeplex.com/	Flickr	Open source -downloads to your computer
ArchiveFacebook 1.4	https://addons.mozilla.org/en-US/firefox/addon/archivefacebook/	Facebook	Firefox add-on
Facebook's 'Download My Information' option	Go to your Facebook Account settings to access 'Download My Information'	Facebook	Will download a zip file that can be opened in .html
LinkedIn – has a built-in way to export contacts	http://www.linkedin.com/addressBookExport	LinkedIn	Offered by LinkedIn
Hanzo Archives	http://www.hanzoarchives.com/	Web and social media archiving	Fee-based native format client-side archiving that claims to handle dynamic web - javascript and flash

PageFreezer	http://pagefreezer.com/	Web archiving and social media	Fee-based automated snapshots with search and timeline functionality
Cloud Preservation	http://www.cloudpreservation.com/	Web archiving and social media	Archives and indexes data from websites, blogs, Twitter and Facebook – 2 week free trial; monthly pricing plans
SocialSafe	http://socialsafe.net/	Backs up Facebook timeline, Facebook Pages, Twitter, LinkedIn, Instagram, Google+ and Viadeo profiles	Created in the UK - Stores locally on your computer; annual pricing plans
Backupify	https://www.backupify.com/	Backs up Google apps, personal apps such as Flickr, Facebook, Twitter and more.	Backs up daily, stores in the cloud; available for PDF download. Free 15 day trial, monthly pricing plans.
ThinkUp	http://thinkupapp.com/	Captures activity on Twitter, Facebook, Google+, and Foursquare	Free open source web application used as an analytic and visualization tool; has ability to capture and store social activity; requires a web server that can run a PHP application
HTTrack	http://www.httrack.com/	Website copier	Free, open source offline browser utility that can be used to create mirror websites with internal functionality
Archive-It	http://www.archive-it.org/	Web and social media archiving	Fee-based service for institutions, has been used by Senators' offices
Heritrix	http://crawler.archive.org/index.html	Web crawler	Internet Archive's open source tool
Contextminer	http://www.contextminer.org/	Web crawler	Open source project that captures contextual information through web crawls. While not an archiving tool persay, this might be helpful in tracking data related to a Senator or committee