

Preserving Social Media

Digital Preservation Topical Note 8



What is social media and how is it used?

Social media has grown into an important channel for organisations to communicate with the public. While websites provide a quick and useful way to publish information, social media allows organisations to interact with the public. Social media comes in many shapes and sizes with a range of different purposes and functionalities. Blogs allow organisations to share substantial information through a more informal platform. Social networking sites such as Facebook and Twitter allow organisations to interact with users and receive feedback. Content sharing sites such as YouTube and Flickr allow organisations to post and organise images and video. These interactions can create valuable records; however, it is not enough to simply capture content. To successfully archive social media for re-use, these data must be indexed and stored in a searchable environment. More importantly, while social media users already generate data at an

unprecedented scale, they are growing exponentially. Organisations should begin archiving social media sooner rather than later as it will only get more difficult and expensive down the road.

Why should an organisation preserve social media?

Social media, like other web content, represents official communication from an organisation. Records created through social media are subject to the same compliance regulations as web content and other electronic documents. As such, social media records are as much a part of an organisation's information and records management as email or reports. Social media also contribute to corporate memory and can provide a rich and vivid picture of an organisation at a particular point in time. If archived social media is well-indexed and searchable, it will be a valuable resource in the future for research and development.

Key Term: Application

Programming Interface (API)

Provides an interface for communicating with the back end of a social media system and enabling such functions as querying or requesting copies of social media data

What are the challenges to preserving social media?

Social media services have characteristics that make them difficult to archive in the same ways as traditional websites. On one hand, the interactive nature of social media makes it difficult to define the content that needs to be preserved. For example, if an organisation wants to archive Twitter data, does it need to capture only tweets published by the organisation? Does it need responses or mentions by other users? On the other hand, the functionality of social media – including the use of embedded images, videos, and URLs – make it very difficult for traditional web archiving tools, like crawlers, to capture.

What are the legal issues that affect preserving social media?

While laws and compliance regulations increasingly demand that organisations archive social media, social media platforms limit and restrict how these data can be used. Platforms control content created by users of their services through a set of terms and conditions. Terms of Service, in particular, dictate what social media content can be preserved and how. Platforms alter their terms and conditions fairly frequently – sometimes as often as every few months. To avoid accidental violation, organisations should schedule set reviews. Remember, national legislation and regulations supersede platform terms and conditions. When in doubt, consult records management and archives staff.



Does preserving social media violate user privacy or data protection regulations?

While social media data, especially in large quantities, holds the potential to reveal personal identities (sometimes of vulnerable individuals), there are ways to ensure privacy. Anonymisation and de-identification are two methods that can help reduce the risk of accidental disclosure. Secure storage of archived social media and limited permissions for access within the organisation will also ensure data protection regulations and individual privacy are protected. Remember this data is generated by the public so it is important to demonstrate that you have their privacy and rights in mind.

How do you archive social media?

The methods for archiving social media depend heavily on how an organisation uses social media and on what platforms. While the records management and archives staff are ultimately responsible for developing a strategy, effectively implementing an effective plan for preserving these valuable records requires support from all staff involved in the use of organisational social media.

Owned Accounts

For corporate accounts for which an organisation is the owner (and holder of log-in information), the ‘download your archive’ function provided by some platforms may fulfil a baseline requirement. Twitter, Google services, and Facebook, for example, all allow account owners to download their own archived content.



APIs to Capture Wider Data

To archive a larger collection of social media, the best and most authentic method of capture is through a platform Application Programming Interfaces.

Purchase Historical Data

In some cases, it might be necessary to purchase historical data (as old as a few months or even a few days) from a platform or from a data reseller. Setting up an API-based collection programme as soon as possible will reduce the cost of purchasing historical data in the future.

For more information on Digital Preservation visit the DPC Website: <https://www.dpconline.org>