Social Media Archive

Go to the Social Media Archive

Collecting social media

The Library collects material to provide a documentary record of life in New South Wales [1] from the earliest times to the present and to provide information services to the people of New South Wales.

The documentary record of life in New South Wales is now, more than ever, digital.

For researchers of the future pondering the question of what life was like in New South Wales in 2017, answers will be born-digital as well as included in books, newspapers, manuscripts and other formats. The Library has developed a Digital Collecting Strategy (PDF) [2] to ensure that we are collecting both digital and analog formats.

Social Media Archive

A significant activity in the Library's Digital Collecting Strategy [1] is a partnership with CSIRO Data 61 [3] to collect and archive publicly-available social media. The Library has been working with CSIRO and their tool Vizie since 2012. Over this period the Library has collected tens of millions of posts.

How does the archiving tool work?

The Library identifies keywords and hashtags which we would like to collect. The Archive then gathers the social media posts that contain these keywords and hashtags from a range of sources, including Facebook, Twitter, and Instagram. The Library has a set of keywords and hashtags that are regularly collected.

We are also able to respond to new issues and events quickly by adding new hashtags which can be used immediately pick up new conversations or conversations from earlier posts that are just emerging as significant.

To address potential privacy concerns the social media archive is designed to only show aggregate data. It allows users to view trends in discussion in social media but does not support viewing content at the level of individual posts.

Aggregate figures are only produced when the number of social media items matching a set of conditions is above a threshold. If the number of items matching the conditions is below this threshold then no aggregate counts will be provided.

The collected data is currently archived with the CSIRO but in the future this data will be migrated across to the Library's digital repository.

The user interface was designed by the CSIRO Data61 [3] team in collaboration with the State Library. The design allows for the data to be displayed geographically, as well as visualised through a 'wheel of emotions'.

What has been collected?

In the year 2016/17, the Library collected 15,700,539 posts. The posts document a range of subjects including politics, topical issues, business, sport, arts, the environment and major events.

Significant hashtags collected include #auspol for Australian political discussion, #nswpol for New South Wales political discussion, and #ausbiz for Australian business discussion.
Significant events documented included:

- the August 2016 Census - "What do you mean they all logged on at the same time?" #census2016 #CensusFail,
- the January resignation of Premier Mike Baird - Thank you, NSW!
- Australia Day - I say it every year but can we just have Australia day on May 8? May 8? M8! Maaaaaate
- Anzac Day - #IStandWithYassmin

In the same way you can go into a library today and look through newspapers and other documentation from the early twentieth century the State Library's social media archive creates a collection of data that will allow researchers in the future to gain insights into the life and times of people living in New South Wales in the digital age.

**What can I do with the Social Media Archive?**

You can view the latest activity and the emotion clock on the [Live section of archive](#) or search for a specific topic or hashtag in the [Explorer section](#).

You can find out more about how Vizie is designed through the [About section of the archive](#).

**What to do if you are concerned about content included in the live interface?**

This a live feed reflecting real time social media activity in New South Wales that is publicly available. There may be occasions when content is considered to breach copyright or other relevant law or contains information that is culturally sensitive.

If you have concerns over any content please refer to the Library’s [Take-down Requests](#). If you wish to reuse information gathered through this tool you must abide by the terms and conditions of the individual platforms.

**How will the Social Media Archive develop in the future?**

An API which will enable researchers to delve further into the archive is currently under development.
The State Library Of NSW Social Media Archive [10]

Video of The State Library Of NSW Social Media Archive


Links