

Preserving personal social media accounts, now and into the future

By Valerie Love

While every year has historic events and defining moments, 2020 saw more than its fair share in just the first six months alone. From the Australian bushfires, to the global Covid-19 pandemic, to the fight for racial justice in the United States and solidarity movements around the world, and the 2020 General Election here, New Zealanders have been experiencing and responding to the global events of 2020 in a myriad of ways.

This article walks you through this process and provides a framework for saving ephemeral digital memories shared via social media. It also provides a list of things to consider and be aware of, including legal regulations, as well as giving an overview of the practice of collecting Facebook content at the Alexander Turnbull Library.

In our digital worlds, we are all inter-connected. Much of the experience of 2020 has been documented online via social media platforms, such as Twitter, Facebook, and Instagram. In recent years, we have seen how global social media advocacy campaigns, like #MeToo, #ProtectIhumātao, #BlackLivesMatter, and #ArmsDownNZ are creating tangible change in our society. People across Aotearoa New Zealand use social media to document the moment, posting photographs from their daily lives, creating videos during noho rāhui/lockdown, and sharing vital information from doctors and experts on how to stay safe. Social media also helped us to connect and understand how the events of 2020 have impacted different communities within Aotearoa New Zealand whose experiences may not have matched mainstream lockdown narratives. The multiplicity of stories and perspectives from this time will hopefully help future generations understand 2020, and see our triumphs, understand our fears, and pay tribute to the lives we have lost.

But with so many of these experiences being recorded and captured digitally, there is a risk that they may not last after all. The world moves

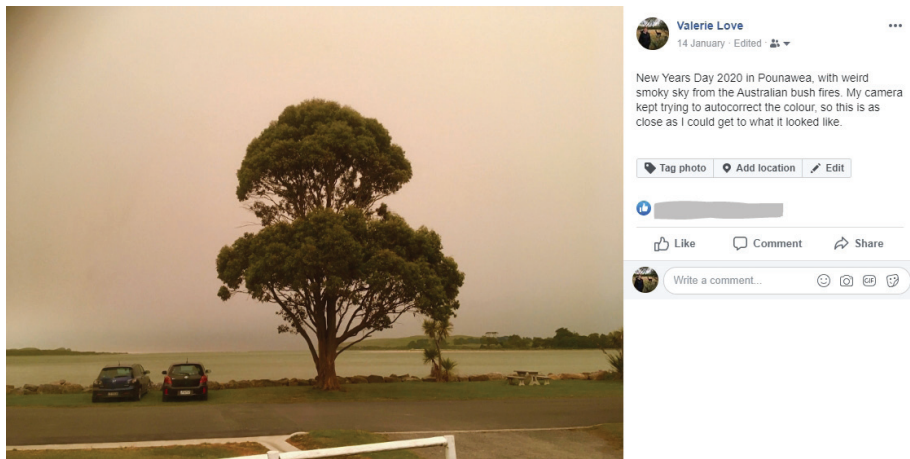


Figure 1: Screenshot from the author’s Facebook account, showing a photograph taken on 1 January 2020 at Pounaweia in the Catlins. The photograph shows smoke visible in Aotearoa New Zealand from the 2019-2020 Australian bushfires. Image supplied by author

fast online, and information comes and goes. Posting content online does not mean that it is actually preserved. Digital files are inherently fragile. They depend on technology, hardware, software, infrastructure, and require people to maintain them over time. Just because Twitter and Facebook exist today, this does not mean they will always be accessible or manage content in the same way in the future. For example, Bebo, Old Friends, Friendster, Vine, and dozens of other platforms where people previously shared and saved content are no longer in existence. It’s therefore important to understand and document social media materials and other digital files, keep track of where they are stored, how they are organised, and ensure that they are regularly checked and backed up to preserve them over time. The GLAM sector has an important role to play in this work through collecting decisions, as well as outreach, guidance, and support¹.

Libraries, archives, and museums across Aotearoa New Zealand currently hold millions of digital files in their collections. However, the vast majority of digital files created will be maintained over time by those who created

them, rather than by a collecting institution. So, if it is important to you and your whānau/family that your stories are preserved and made available into the future, there are practical steps you can take now (and encourage others to as well) in order to preserve Aotearoa New Zealand's digital history.

Social media as an archive - some considerations

Within my own social media bubble, I have noticed a surge of people taking photographs at events and posting them to Facebook groups or on Twitter as a way of making them publicly available. This can be a highly effective way of sharing and amplifying the reach of content. However, if you choose to use social media as a personal digital archive, there are a few considerations to be aware of.

Firstly, data sovereignty can be an issue with overseas platforms, particularly for Māori digital materials. Data sovereignty generally refers to understanding that data is subject to the laws of the country in which it is stored. Māori data sovereignty recognises that Māori data should be subject to Māori governance, and ensures that data for and about Māori can be safeguarded and protected.²

Many cloud storage providers host their data on servers overseas, so if the location of your data is important to you, make sure you are aware of where your data will be held, and what protections are in place. Instagram's terms of service, for example, state that it does not take ownership of its users' content. However, the act of posting content to Instagram grants the platform a "non-exclusive, royalty-free, transferable, sub-licensable, worldwide license to host, use, distribute, modify, run, copy, publicly perform or display, translate, and create derivative works of your content (consistent with your privacy and application settings)."³ While the person posting technically still owns the content they post, Instagram has many of the same rights as the original creator of the content.

Secondly, it is important to be mindful that there is a loss of control over original files when using social media as a *de facto* archive. Other users can screenshot, download, or otherwise repurpose materials in ways the creator may not have originally intended. While social media is designed to be social, and encourage interaction and engagement with others, when an image goes viral it can be easy for the person who posted it to

become overwhelmed by comments (both positive or negative).⁴ Depositing contemporary digital materials in a library or archive where they can be catalogued with copyright information and explicit statements guiding use and reuse can help to protect material from misuse, and establish boundaries, although this option may not be suitable for all users.

A third consideration is that most social media platforms compress image and video files, so what is posted online and what you can download back from the platform, by requesting your data, may be lower quality than the original files. This may not matter for some content, but it is something to be aware of. For this reason, uploading files to a social media platform should not be considered a replacement for keeping digital material of personal significance.

And finally, social media platforms are ultimately commercial data enterprises. Their business model is not designed to preserve content in perpetuity. As the Facebook news ban in Australia in February 2021 showed, platforms are able to – and do – remove user content at will.⁵ This could have catastrophic consequences for historians of early 21st century life if social media material is no longer available.

Collecting Facebook at the Alexander Turnbull Library

The Alexander Turnbull Library first began collecting Twitter content in November 2016, with harvests of publicly available Twitter content related to the 7.8 magnitude Kaikoura Earthquake.⁶ Since then, the Library's web archivists have conducted regular Twitter crawls on significant political and cultural events in Aotearoa New Zealand using Documenting the Now's Twarc command line tool.⁷ Twarc allows users to archive Twitter by querying the Twitter API and returning each tweet in the query as a JSON object.

In 2018, the Library began to consider ways of collecting Facebook content as well. With nearly 3 million active users in Aotearoa New Zealand, for many people, Facebook is the digital equivalent of scrapbooks and photo albums. It is where people react to news and connect with family and friends. Whether or not people think of it as such, it is an archive of our daily lives. As archivists, we know that these materials will be of interest to future researchers and family historians, and that we need to be collecting this content now, so it does not just disappear.

The Alexander Turnbull Library launched a social media archive project in August 2019, beginning with a public call for New Zealanders to download and donate their Facebook account archives to the Library. The project was part of the Alexander Turnbull Library's centenary celebrations, as which reflect on the Library's first 100 years of collecting, and consider what the emerging and anticipated research needs of the future might be.

The Facebook Archive Project instigated by the Library asked people to download the archives of their personal Facebook account and donate the content to the Alexander Turnbull Library.⁸ Our aim was to collect a set of Facebook archives to create a diverse snapshot of born-digital social life of the early 21st century. We launched the project via a blog post on the Library's website, as well as a social media campaign on Facebook (of course) and Twitter.⁹ There was also an article about the project on news website Stuff.co.nz.¹⁰

To make it as easy as possible for people to donate their Facebook account archives, we created a form¹¹ on the Library's website. This allowed participants to provide information about their Facebook account, select the conditions for access (either currently available onsite in the reading room, or closed until a future date), and tell us about their account and why they wanted to donate it. We worked with the Department of Internal Affairs' (the Library's umbrella organisation) Legal Services team to draft the terms of donation and included them in the online form to ensure donors were properly informed of what we would be collecting. This also established a clear chain of custody between the donor and Library for the material. After filling out the form, users were guided to a Dropbox link where they could submit the account archives as a zip file. We also created a brief online video outlining the downloading and submission process.¹²

We were unsure of what the response and rate of participation would be like. As it turned out, we got good feedback about the project and the fact that we were doing it, but there was less appetite for participation than we had hoped. We only received about 20 Facebook account archives, rather than the 50 – 100 we had originally aimed for. However, they cover a range of geographic locations and interests; everything from organic farming, sourdough making, queer life, Steampunk, New Zealand politics, and more.

Once the files were received by the Library, they went through the usual appraisal processes for all incoming born-digital material. The technical checks confirmed that the zip files were valid and contained the required files. If not, we sometimes had to contact the account owner and ask them to resubmit the files. We also confirmed that the content was as expected, and that the donations met the criteria; i.e. that the donor is the copyright holder/owner of the account, the content has a relationship to Aotearoa New Zealand and/or the Pacific, and that they give the Library permission to archive their Facebook content in the National Digital Heritage Archive (NDHA).

We know that people's Facebook archives are very likely to have personal and private information, not just about them, but also about their friends and whānau. As such, none of the Facebook archives will be openly accessible, even if the donor's own account is currently public. The web form allows donors to choose a level of access restriction for their archives – either available onsite only in the Katherine Mansfield Reading Room or restricted for 25 or 100 years. Most donors were happy to have their account information available to researchers onsite without further time restrictions.

Download an archive of your social media accounts

With the passage of the General Data Protection Regulation (GDPR) in Europe in 2018, tech companies began to improve the processes for individuals to access their personal data.¹³ These processes do change from time to time, so check the platform's help pages for the latest information and instructions. With any platform, it is good practice to check your data regularly, and see what personal information about you is being stored. If your social media archives are important to you as a record of your life, perhaps set up a regular schedule for downloading them every six months or once a year.

The following sections outline how to download content from various social media accounts. Please note that the information and instructions for downloading data from social media platforms were correct at the time of writing. However, platforms do change their interfaces and offerings regularly, so it is best to read the instructions offered on individual sites for further information.

Facebook

Facebook allows you to download your account archive as a zip file in either JSON or HTML format. The JavaScript Object Notation (JSON) format is machine-readable, and this allows it to be more easily processed or uploaded to another system. Hyper Text Markup Language (HTML) is machine-readable as well, but offers a more user-friendly data structure, including an **index.html** file, which allows you to look through your archive using a web browser. Facebook also offers the option of choosing the media quality of the download – for archival purposes, select “High.” This means the download may take longer, but your images and videos will be of better quality. You can choose exactly what information you would like to include or exclude from the download, such as messages, activity in groups, and more.

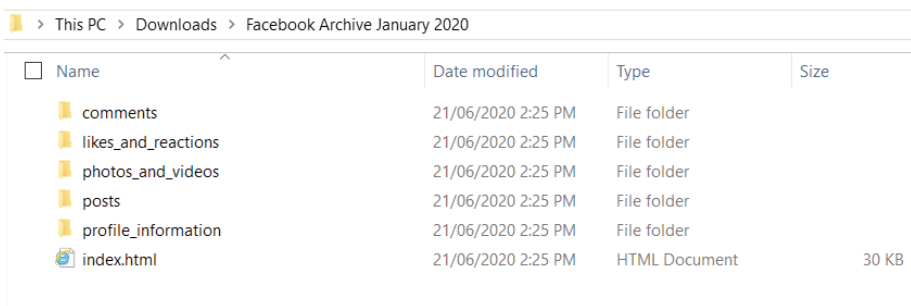


Figure 2: Screenshot of Facebook Archive data. This is quite a simple one where just a small amount of the available data has been downloaded. Image supplied by author

If you manage a Facebook Page, you can download that as well. However, Facebook does not offer the ability to download the archive of a Facebook Group. Any posts that you have made to a Group will be part of your own archive, but it is not currently possible to download the Group archive as a whole entity. So, if you have a community group with photographs or other materials that you wish to preserve, you may need to use other web archiving tools, such as the Webrecorder desktop app.¹⁴

Flickr

Flickr offers several options for downloading your data and photographs. You can download either your entire camera roll as a zip file, or select individual albums to download, which will maintain the photograph's database number, name, and any EXIF (Exchangeable Image File Format) metadata automatically embedded into an image file by the digital camera it was taken with.¹⁵ Alternately, you can request your full Flickr data, which will supply zip files with your photographs and videos at full resolution, as well as JSON metadata files with your account information.

Google Takeout (YouTube, Gmail, Hangouts, Blogger, Google Photos and more)

Google products, such as YouTube, Gmail, Hangouts, Blogger, Google Drive, Google Photos, Google Maps locations, Google+ and other products, can be downloaded via the Google Takeout service. If you have an email account that you care about, it is a good idea to regularly download an archive of your account as either .PST (Personal Storage Table, used by Microsoft email products) or email mailbox .MBOX files.

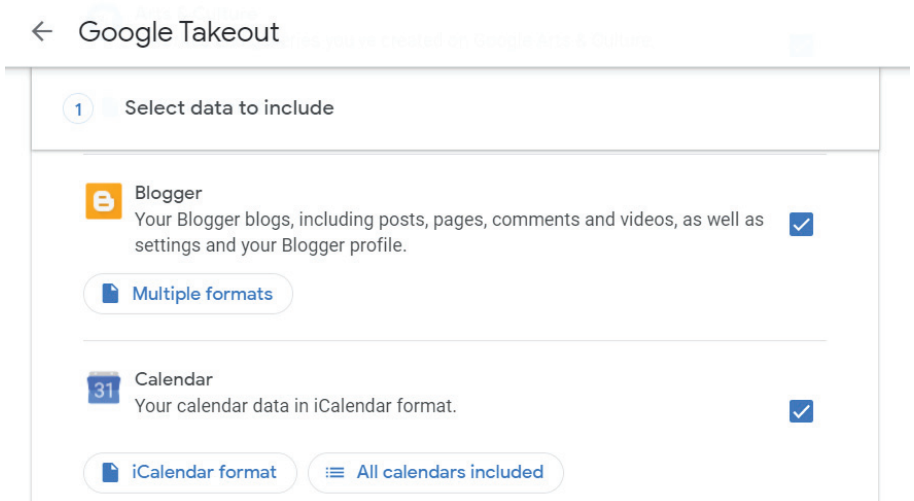


Figure 3: Some of the data you can download through the Google Takeout interface, <https://takeout.google.com/>. Image supplied by author

To access Google Takeout, simply go to your Google Account settings and click 'Data & Personalisation'. Scroll down to the 'Download your data' link, then select the data you wish to download. Unlike most archives which limit the frequency of your downloads, you can set up a one-time archive, or have Google schedule regular data exports for you.¹⁶

Instagram

The process for downloading your Instagram data is similar to that of its parent company, Facebook; however, unlike downloading from Facebook, Instagram data is provided in a standard way without any options to choose from. The Instagram archive is provided in JSON format, and includes posted videos, comments on photographs, likes, searches, messages and more.¹⁷ The Instagram download option is available under 'Privacy and Security', and you can either view your data, or request a download to be emailed to you.

LinkedIn

LinkedIn offers the ability to customise the data files you are most interested in from your download, to make it easy to find articles, contacts, profile information, or other data you may have shared over the years. From the top menu bar, select the dropdown arrow beside your profile photo where it says 'Me', and select 'Settings and Privacy'. Scroll down to the section titled, 'How LinkedIn uses your data' and click on 'Get a copy of your data'. You can then either download your full archive, or specific data files from it.

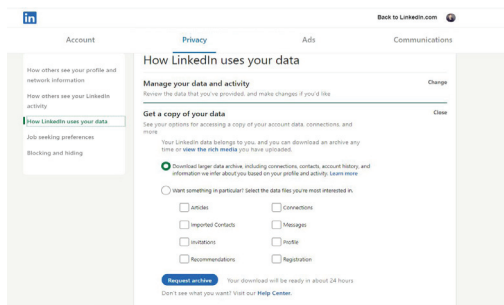


Figure 4: LinkedIn data download options. Image supplied by author

Pinterest

Unfortunately, Pinterest does not offer a way of downloading your boards. However, if you are mainly interested in having a visual representation of your boards, you can use a tool like Conifer (previously Webrecorder.io) to create a web archive.¹⁸

If your board is small, you can use your browser's 'Save as' functionality to save a copy of the webpage as a **complete html** object, which will save a visual representation of the site, as well as create a folder of images. However, the pins on Pinterest boards load as you scroll, and your browser's web archive tool will only save the portion of content currently loaded on the page. To fully archive a Pinterest board with a lot of pins, you may need to scroll down and 'Save as' multiple times.

Reddit

Reddit doesn't have a download function to access your data as a whole, but there are a series of links in their Help section where you can access various types of data, such as posts you have made, and communities you moderate.¹⁹ To obtain a copy of your full data, you need to fill out a data request form, though it may take up to 30 days for Reddit to prepare your data.

Snapchat

Snapchat is designed to be ephemeral, where the images shared disappear after 24 hours, so if you have not opted to save them at the time of sharing your snap, you cannot get them back. However, you can download your account information, your profile, public Snapchat stories you've contributed to, and your Snapchat history (the latter lists to whom and when you sent snaps).²⁰

TikTok

TikTok, the second most downloaded app in 2019, is known for making it easy to download and save other users' videos.²¹ Unfortunately it does not offer the same ability to download your own archive. Currently the app does not provide a way for users to automatically request or download their data.

If you use TikTok and value the content that you post there, make sure that you backup copies of your videos somewhere else as well.

Twitch

Twitch offers users the option of automatically storing content at the time of posting, but depending on your membership level this storage may be for as short as 14 days for a basic user, or 60 days for prime users.²² Again, the best option is to make sure that you have saved a backup copy of your videos, or download them from Twitch on a regular basis.

Twitter

Twitter allows you to download your complete account archive. If your Twitter account is connected to a Periscope account, you can download that data too. The file downloads as a zip file containing various folders of data, including a **tweets.csv** file that contains all of your tweet text and metadata formatted into a spreadsheet. The Twitter download also includes an **index.html** file that allows you to browse tweets by month using a basic web interface. The download includes account information formatted into JSON and downloads of images and other media posted to the account.

Vimeo

With Vimeo, you can download your videos individually, but there is no easy way to download an entire channel at once, unless you make use of scripting tools.²³ Paid members do have the ability to store their original, unencoded source files on Vimeo as a form of file storage.

WhatsApp

Also owned by Facebook, WhatsApp is the most popular messaging app in the world, with over 2 billion users.²⁴ Users can request their data in the Settings menu, which includes settings, profile information, and group names. However, be aware that requesting your data does not

actually export your chat history. Instead, you will need to export each chat individually. Exact instructions vary depending on whether you are exporting from an Android or OS operating system.

So, you have downloaded your social media data. Now what?

Now that you have downloaded your social media archives, designate a secure place to save them. This might be in a folder on your computer or in the cloud, with backup copies on an external hard drive, for example. The zip file downloads will often be labelled by default with just a series of numbers and letters, so make sure to give them meaningful names, such as *Your_Name_Instagram_archive_June_2020*.

If you are considering collecting social media archives, ensure that you ask donors for as much contextual information relating to the files as possible, or what they think is noteworthy about their social media archives. Remember that social media archives contain more than just our own personal data - they often relate to friends, whānau, and extended social networks. It is important to have robust policies in place in terms of access, use, and/or takedown, if necessary.

As Jessica Moran has written:

For archivists, social media platforms have the potential to “produce and disseminate a record of our cultural heritage” (Social Feed Manager, 2017) and can be the site of valuable records and documentation. But with what can sometimes feel like the exponential growth of social media platforms, users, and content, the challenge for archivists is both deciding what to collect and how to collect. This is both a technical challenge, and an archival appraisal challenge. It is a technical challenge in that different platforms will sit on top of specific technology archivists need to have at least a basic understanding of, but also that technology is constantly being adapted and updated. It is an appraisal challenge because we cannot and should not be collecting everything. Part of our role must also be to make sense of what records or content will have long-term value in their particular cultural and organizational context.²⁵

As archivists we must be willing to take up these challenges. We need to experiment, to learn, and to know that as platforms change, and the ways in which people create and share information changes, the way we collect and document society must change as well. But more often it is our role not just to collect, but to ensure that others have the information and tools they need to manage their own digital files over time. Digital preservation is a challenge that we all must face together.

Endnotes

1. Galleries, Libraries, Archives and Museum
2. See also "Principles of Māori Data Sovereignty. Brief #1 October 2018. Te Mana Raraunga Māori Data Sovereignty Network. Accessed 3 March 2021. <https://cdn.auckland.ac.nz/assets/psych/about/our-research/documents/TMR%2BM%C4%81ori%2BData%2BSovereignty%2BPrinciples%2BOct%2B2018.pdf>
3. Instagram. "Terms of use". Accessed 24 June 2020. <https://help.instagram.com/478745558852511>
4. See, for example, Pruett, Sarah. "The Day My Two-Month-Old and Husband Blew Up the Internet." Medium (blog), January 12, 2019. Accessed 24 June 2020. <https://medium.com/@sarahpruett/the-day-my-two-month-old-and-husband-blew-up-the-internet-37915e65b9ba>
5. Ryan, Siobhan. "Facebook's Australian news wipeout showed it can delete our history at any time." Accessed 3 March 2021. <https://www.theguardian.com/commentisfree/2021/mar/01/facebook-australian-news-wipeout-showed-it-can-delete-our-history-at-any-time>
6. National Library of New Zealand. "Kaikoura earthquake: collecting social media". Accessed 24 June 2020 <https://natlib.govt.nz/blog/posts/kaikoura-earthquake-collecting-social-media>
7. Github. "DocNow/twarc". Accessed 24 June 2020. <https://github.com/DocNow/twarc>
8. National Library of New Zealand. "Facebook Archive Project". Accessed 24 June 2020. <https://natlib.govt.nz/collections/donations/donating-digital-items/facebook-archive-project>
9. National Library of New Zealand. "Is your Facebook account an archive of the future?" Accessed 24 June 2020. <https://natlib.govt.nz/blog/posts/is-your-facebook-account-an-archive-of-the-future>
10. Deguara, Brittney. "National Library creates Facebook time capsule to document New Zealand's history." Accessed 24 June 2020. Stuff. September 5, 2019. <https://www.stuff.co.nz/national/115494638/national-library-creates-facebook-time-capsule-to-document-new-zealands-history>.

11. National Library of New Zealand. "Step 3 — Facebook archive donation form." Accessed 24 June 2020. <https://natlib.govt.nz/collections/donations/donating-digital-items/facebook-archive-project/step-3-facebook-archive-donation-form>.
12. National Library of New Zealand. "Step 2 — Download your Facebook Archive". Accessed 24 June 2020. <https://natlib.govt.nz/collections/donations/donating-digital-items/facebook-archive-project/step-2-download-your-facebook-archive>.
13. Binder, Matt. "TikTok got an 'F' in our data accessibility rankings. Facebook got an 'A.'" MashableAustralia. Accessed 24 June 2020. <https://mashable.com/article/social-media-data-request-report-card/>
14. Conifer Webrecorder desktop application. Accessed 3 March 2021. <https://conifer.rhizome.org/>
15. Dove, Jackie. "How to Download Your Photos From Flickr Before They Disappear". Tom's Guide. Accessed 24 June 2020. <https://www.tomsguide.com/us/flickr-photos-download-how-to,news-28932.html>
16. Binder, Matt. "TikTok got an 'F' in our data accessibility rankings. Facebook got an 'A.'" MashableAustralia. Accessed 24 June 2020. <https://mashable.com/article/social-media-data-request-report-card/>
17. Burgess, Brian. "How to Download a Copy of Your Instagram Data." groovyPost (blog). December 31, 2018. <https://www.groovypost.com/howto/download-archive-all-your-instagram-data/>
18. See <https://conifer.rhizome.org/>. Accessed 11 November 2020.
19. "Where and how can I access my Reddit data and information?" Reddit. Accessed 3 March 2021. <https://www.reddithelp.com/hc/en-us/articles/360043483511>
20. Rutnik, Mitja. "How to save Snapchat videos on Android". Android Authority. Accessed 24 June 2020. <https://www.androidauthority.com/how-to-save-snapchat-videos-890021>
21. Sheng, Wei. "Tiktok is the target of yet another data privacy probe". TechNode. Accessed 24 June 2020. <https://technode.com/2020/05/09/tiktok-is-the-target-of-yet-another-data-privacy-probe/>
22. "How To Archive Your Broadcasts In Twitch". Techjunkie. Accessed 24 June 2020. <https://www.techjunkie.com/archive-broadcasts-twitch/>
23. See: <https://github.com/ytdl-org/youtube-dl> for example
24. Fish, Tom. "WhatsApp Export Chat: What does export chat mean? How to export conversations on WhatsApp". Express. Accessed 24 June 2020. <https://www.express.co.uk/life-style/science-technology/1140438/whatsapp-export-chat-meaning-how-to-export-whatsapp-conversations-iphone-android>
25. Moran, Jessica. "Social media archiving", In Handbook of Archival Practice. Rowman & Littlefield, forthcoming.