Somewhere between automation and the handmade

Interview with Rosemary Grennan

Leftovers is a project that seeks to create a shared online archive of radical, anti-oppressive, and working class movements, and the material traces they have left. The platform aids the dissemination of archived ephemera from these movements, campaigns, and struggles, casting light on histories of resistance from below by opening up archives of radical dissent. Leftovers consists of a website (https://dev.leftove.rs) and an archive backend (https://archive.leftove.rs).

Everyday Technology Press invited Rosemary Grennan for an interview to speak about the Leftovers archive (https://leftove.rs), the range of materials that it holds, and the ways in which MayDay Rooms worked on structuring it all.

Could you introduce the Leftovers archive?

The project is a collaboration between MayDay Rooms in London and 0x2620 in Berlin. Currently we have nearly 18,000 items, which are scans of original documents, nearly all of which have been OCRed (optical character recognition), meaning that all documents are fully searchable. The kind of material we hold covers all types of political ephemera from pam-

phlets, posters, and photographs. Although MayDay Rooms initiated the project, Leftovers does not solely consist of our digital collection; it draws in digitisations from many different sources: from torrent files of 1970s newspapers, to an autonomously-run online collection of the Ultra-Left in France, to Women's Liberation movement material from state archives.

At MayDay Rooms we try to rethink how we work on, and with, archives of struggles, and see our collections as something that should be active in the present through strategies of open access, activation, reuse, and through building the archive as a collective resource. This ethos is something that we have attempted to bring to how we approach our digital archives as well as the digitisation of our collections.

How is the Leftovers archive structured at the moment?

Before going into more detail about the structure of the archive, I wanted to give a little preamble about the thinking behind, and the motivations for, creating Leftovers, as this heavily influenced the structure the collection now takes.

There is quite an emphasis in contemporary archival practices on the digitisation of collections for reasons of both access and preservation. Despite these intentions, access is often still restricted by questions of rights and digital preservation takes a lot of resources (storage etc) that smaller independent archives do not have.

The digital collections that result from archival digitisation projects often simply mirror a physical collection and are there to embellish the catalogue. At MayDay Rooms we wanted to think about what our digital collection could do that is different from our paper collection, and how our guiding principles can influence a digital formation. So we started to think about how digital possibilities of the archive can bypass traditional concerns of preservation in favour of dissemination and truly open access.

We have been inspired a lot by the work of friends and comrades from different "shadow libraries," particularly Aaaaarg, and Memory of the World. There was an initial idea that MayDay Rooms' digitisations form part of these two libraries' collections, and we still host selected material on these sites. However, we soon realised that historical ephemera (posters, pamphlets, flyers, bulletins etc) not only require very particular attention but also present interesting possibilities in regards to a digital archive.

This is also why we chose the name Leftovers; as well as being a joke, it is also a comment on the nature of political ephemera. The book is in some ways a relatively self-contained durable object – it has a blurb, a recognisable author – that doesn't need additional material to become understandable, and its use in some ways is predetermined. Whereas po-

litical ephemera has a different temporal scope, it was not meant to endure, and its contemporary use is different from its first production. Its original intention. The leaflet to mobilise for a protest, the bulletin that communicates actions on a picket line, or newspapers that maintain organisational forms are meant to mobilise quickly and communicate in the moment but not to last. For these reasons ephemera is non-authoritative, and represent fragments of historical moments through different tendencies, so it needs a critical mass of comradely material and different archival strategies to make it understandable.

The structure of our digital collection has tried to reflect this structure, using flat relations between objects rather than hierarchical ones, and developing our metadata categories instead of using inherited conventions. Our collaboration with 0x2620 meant we could further experiment around the different processes and relations a digital archive can facilitate. Although the software Pan.do/ra was originally developed for video, we have worked with Jan Gerber at 0x2620 for the last few years to see if the way in which the software decomposes videos and makes each frame accessible could do the same for digital documents.

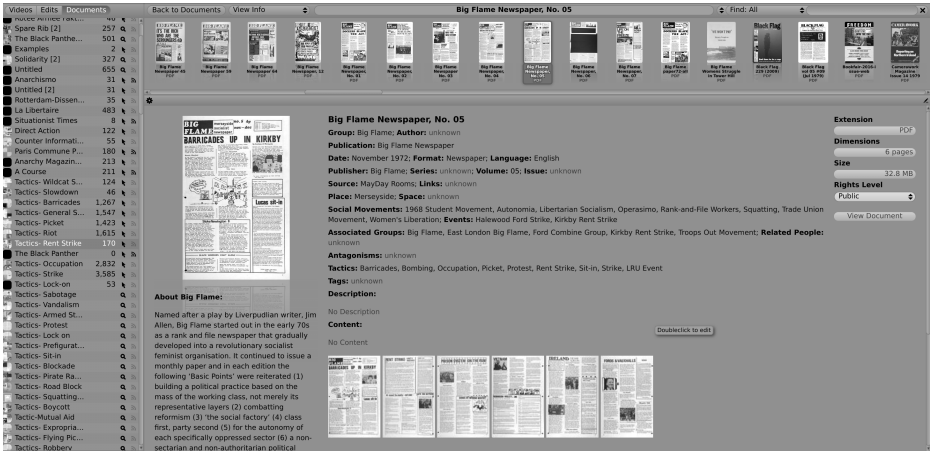
Leftover does not only represent MayDay Rooms' digital collection but is rather an "archive of archives" where we have pulled together existing online repositories and resources from all the types of institutes, collections, files, and folders into one platform. At present, materials in the archive come from many different sources and are not usually the only copy of that scan. The metadata around each object always links back to the source that we got it from. We hold material from some sisterly archives such as the Sparrows' Nest Library and Archive in Nottingham, and other material we found from different corners of the internet (torrents, smaller archival collections, state archives). Some of the processing we have done on the documents, and the functionality 0x2620 developed for the archive, has helped us think of different ways of making connections between documents that come from different collections, countries, tendencies, and groups.

I thought it might be good to go into a bit of detail about how we developed one of our metadata fields, "Tactics." As I said, all the material in our archive is OCRed and there is a full text search functionality so that you can search *within* the document, not only for data *about* the document. This might sound a minor technical point but actually is highly significant in opening up digital archives and using the actual document's content as the basis of classification. In Leftovers you can search for a word or phrase and it will bring up every document that includes it. Through this we made a list of different tactics of left and anarchist movements and searched the documents for them. Some of the results are below with their occurrences

in the archive:

Occupation (2858)
Rent Strike (164)
Riot (1630)
Picket (1451)
Strike/Grève (3640)
Direct Action (2401)
Rent Strike (170)
Prefigurative (18)
Sabotage (2053)
Armed Struggle (343)
Protest (3399)
Sit-in (905)
Blockade (607)
Pirate Radio (76)
Collective (4584)
Road Block (42)
Boss-napping (4)
Wildcat Strike (124)
Squatting/Squat (614)
Slowdown (46)
Boycott (1367)
Theft (609)
Forgery (1)
Barricades (1310)
March (3850)
General Strike (245)

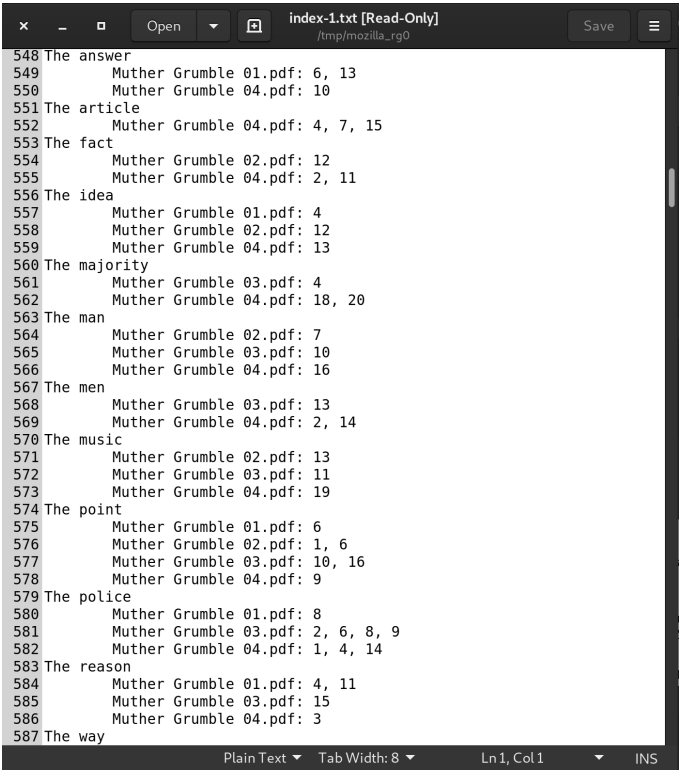
Some of these are terms that occur too many times or are too broad, such as “Occupation,” to be a useful way of filtering an item, but some such as “Rent Strike” are specific enough to be a useful means of linking up documents. We thought that this category of “Tactics” was a useful one to reorientate the collection as something that can be used as resources for current struggles to integrate tactics of the past that might have been forgotten.



In another conversation we had, you mentioned you worked with Natural Language Processing (NLP) tools for this archive and earlier you mentioned that all the documents have been OCRed. So we are wondering how NLP had been applied to the documents? What influenced the conceptualisation of those operations?

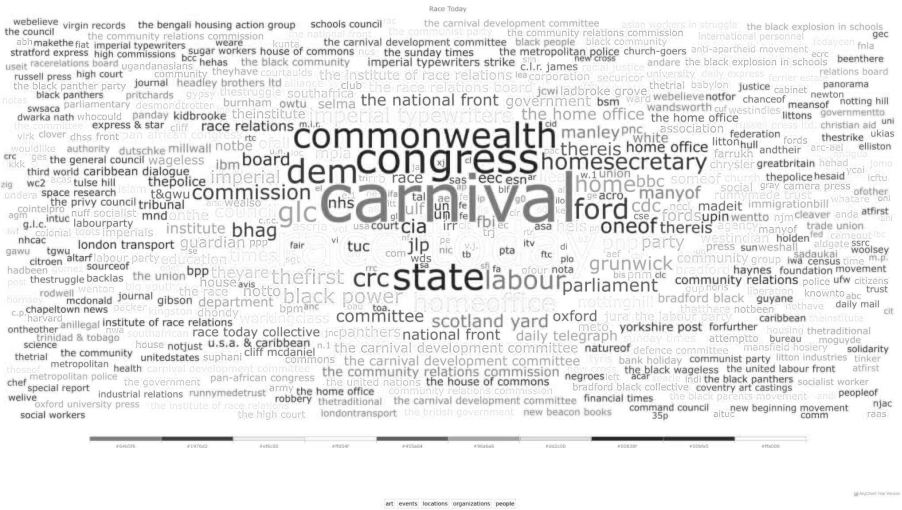
Sean Dockray was the first to use NLP on Leftovers, applying the entity extractor from the spaCy library to help create an index for one of our publications, *Muther Grumble*. He described the process as somewhere between automation and the handmade, and I think that this really characterises the kind of experiments we have been doing with NLP since then. We have used NLP as a research tool to try and get deeper into the documents' content rather than the usual mode of algorithmic analysis, which only looks at the derivatives of the object. The process has definitely brought up more questions around categorisation than answers to it, as it has unearthed many themes, entities, people, and places that we didn't know existed across the collections. From spaCy, we used their libraries to extract arts, events, organisations, and people, and applied these across the entire publication. We are now undergoing a process of sifting through the noise of the results and figuring out what might be relevant to form into different categories, or which names and organisations we should search across the archive.

There have been many different strategies of inputting metadata on Leftovers. Some of this inputting has been automated when data scraping the collection, others have been more of a derive through the collections by those who have knowledge of the material inputting as they go. However, the processing of the catalogue doesn't usually mean the person categoris-



ing the item has fully read the document, and this becomes an increasingly impossible task as the archive grows and grows. So although NLP "reads" the document in a very particular, partial, and biased way, the tension between the actual results of the process and the material in the archive has often prompted us to look for different things in the archive and to read it in multiple ways.

One of the ways of sorting the result of the different NLP scripts was to create word clouds to visually look at the most frequently occurring terms (see above). For example, in the word cloud that showed people, one of the most prominent names was Ronald Reagan, but this data did not fit into any of our categories relating to people, which mainly represent comradely relations. NLP was better at pulling out these known entities like Reagan rather than minor figures in left history, as the libraries have been trained on certain data sets. This could be immensely problematic as a tool of categorisation for an archive of radical ephemera. However, by thinking these relationships through and thinking what to do with this data we came up with the category of "Antagonisms." We thought this showed the archive to be partisan and not a neutral historical collection, as well as establishing



an oppositional relation to some of the data that the NLP was producing.

> You've already mentioned that Leftovers came together as a collaboration with 0x2620. Are there more collaborators involved in the making of this archive?

Leftovers was initiated by myself at MayDay Rooms, and was developed in collaboration with Jan Gerber from 0x2620. Anthony Iles from *Mute Magazine* joined the working group in 2020 and has been very active in finding material and inputting metadata. He also helped produce our first online exhibition based on the digital archive called *Print Subversion in the Wapping Dispute* which can be found here https://exhibitions.maydayrooms.org/wapping.We got a small grant to further develop our interface in 2020, and for this we worked with Gemma Copeland and Robbie Blundell from Evening Class, a design collective in London, to design the front end. You can see the work-in-progress version here, https://dev.leftove.rs.

However the wider question of who contributes to Leftovers is a complicated one, as some people contribute without knowing they do so. For example, *The Black Panther,* the newspaper of the Black Panther Party, was scanned at a university in America and then did the rounds on the internet as a torrent. We downloaded it and OCRed the scans and uploaded it to the collection. We were not the ones who scanned it, downloaded it from the university, and distributed it online and the scans are not only hosted on Leftovers, but we felt it was important that it become part of the archive and be freely accessible. Other contributors are comradely archives such as the Sparrows' Nest Archive and Library (UK),

Archivio Grafton (IT), Rebel.info (IT), and Archives Autonomies (FR) who we have established relationships with through Leftovers. We hope that through pooling and sharing digitisation, Leftovers will help build a network of archives that share documents and resources.

The reason it is not solely a MayDay Rooms project is that eventually we would like Leftovers to become a shared archive which contains many active users, with multiple archives and users uploading and maintaining the material in the collection. We see ourselves as custodians of the material in the collection rather than having some kind of exclusive claim.

> Are there examples where the structure of the archive has been negotiated with others? By others we mean collaborators, audiences, but also perhaps creators of the material in the archive.

Although Leftovers has not been officially launched, it is already being used by smaller archives that have only just started embarking on creating digital collections and want to use Leftovers to host their collections and bring their material into conversation with the rest of the archive. For example, we will be running a workshop in November 2020 with Glasgow Housing Struggles Archive, a new project from members of a tenants' union called Living Rent aimed at uncovering Glasgow's hidden history of squatting, rent strikes, and council tenant organisation, to look at how they can use Leftovers to build a resource and take the archive into everyday organising. We hope that by having nascent archives involved, they will also contribute to the process of building relations between documents and, in turn, help us structure Leftovers in a way that works for all collections.

> Can you say a bit more about the authorship of the material itself? How does the archive relate to different forms of authorship, such as collective or anonymous ones?

That's an interesting question and speaks again to the different kinds of metadata fields we developed for Leftovers. The field of "Author" in the collection is almost completely redundant, as most material in the archive is produced by a group, a collective, or is anonymous. This is an indication of what kind of material the collection holds.

Some material is intentionally authored under a group name. In some of these cases the writing of newsletters, pamphlets, positioning papers, and bulletins becomes a form of internally constituting groups, and here writing and action combine as a form of political organisation. Other material remains anonymous or produced under a group name not as choice but as a societal position in relation to the state, whilst other groups are

actually a single person masquerading as a group in order to explore imaginary formations. Other group names are actually descriptive tactics such as “The Angry Brigade.” With all of these examples anonymity is not merely dropping one’s name, but speaks to the complex nature of the document’s production. By fore-fronting these relations in thinking through the structure of our collection, we hope that the use of relational metadata categories around production and social movement contexts will help to orientate categorisation away from authorship.

As well as the anonymous or group producer in the archive, there are the invisible figures that are integral to the political movements the collection is part of. As I mentioned before, through using Natural Language Processing to make lists of all the names that occur *in* the material, we were able to find many figures that have been forgotten and excluded in favour of single authors. Although these are currently only indexes and have not been entered into metadata for the material, it is perhaps a good indication of the thousands of people who produce material and often prop up the category of author, or are behind prominent movement figures.

Leftovers is trying to undo proprietal forms of ownership over the archive in favour of an archive that is a common resource and will eventually be owned in common too. The destabilisation of the field of the author, in favour of a form that acknowledges all those who went into the production of the material, is very much part of this process.

How does the archive support dissemination of its material?

https://twitter.com/ArchivioGrafton/status/1357425808768385025?s=20

I really liked this tweet from Archivio Grafton about their material on Leftovers, where they say, “when you publish something on the net, let everyone take it and freely distribute it.” It points to the fact that the aggregation of all this radical ephemera is an act of redistribution itself. We only collect things that were part of a political movement and that were shared publicly at the time they were produced, and believe they were produced in struggle and through this are collectively owned.

MayDay Rooms is dedicated to the collective “activation” of historical material – where we aim to not sit passively on archival “holdings.” With Leftovers we were really interested in developing tools, and ways of disseminating, integrating, and re-using the collection rather than it just being a repository where material is merely stored. I can’t remember who said it, but it’s a sentiment that has influenced our approach to an active archive: “The best way to preserve a film is to project it.” Digitising a document,

making the whole thing freely available, downloadable, and readable is the first step towards dissemination.

One of the publications in Leftovers – *Spare Rib* – was originally digitised by the British Library at huge expense and the copyright was contained through Digital Rights Management. After the UK left the EU, the copyright directive that covered the digitisation no longer applied and the digitisations were taken down. To my knowledge Leftovers now holds the only digital copies of this material, which represents the biggest Women's' Liberation publication in UK history. We can see here the fragility of institutional collections that do not commit to open access. We believe that creating collections where you always have access to the scan or object itself is a good step towards facilitating new forms of distribution and back-up that go beyond the original collection.

There is a text called "HyperReadings" (https://samiz-dat.github.io/hyperreadings) by Sean Dockray, Benjamin Forster, and Public Office, which I think really articulates well this idea of a "libraries of libraries," where items are not confined to a single copy in a single universal library but are partially manifest with many different individuals, groups, and institutions. By breaking down categories and ways of collecting that usually confine a document, and using different processes to make connections between what would previously be atomised material, you can begin to circulate material in different ways. All material in Leftovers has a "Source" field that takes you back to the original source of the digitisation, and we hope that aggregating these materials on one platform not only brings them into proximity with one another but also highlights the work of many small independent archives.

Some of the other strategies that could be described at dissemination at the moment are an exhibition of our digital material that I mentioned before (https://exhibitions.maydayrooms.org) publishing highlights from the archive, a project called MayDay Radio (https://audio.maydayrooms.org), which is an independent collective that use the archive as starting point for audio pieces and experiments. This year we have collaborated with them to do a series of residencies with Leftovers material and the first of these pieces, *Abolitionist A-Z*, will be launched soon.

We have also been working on a collaborative tool for working together on the digital collection and creating new ways of interrogating the material remotely. The tool enables users to take clippings from the documents in the archive, then recombine and annotate them to create "scrapbooks" or montages. We held a series of workshops to test and further develop these tools, and through this co-created different scrapbooks around the topics of Health Autonomy, Abolitionist Struggles, Rent Strikes, and Radical Spaces.

You can see some of the results here: https://dev.leftove.rs/\#scrapbook.