Edited by

Memory of the World

Guerrilla Open Access

Christopher Kelty

Balázs Bodó

Laurie Allen

Published by Post Office Press, Rope Press and Memory of the World. Coventry, 2018.

© Memory of the World, papers by respective Authors.

Freely available at:
http://radicaloa.co.uk/conferences/ROA2

This is an open access pamphlet, licensed under a Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

Read more about the license at:
https://creativecommons.org/licenses/by-sa/4.0/

Figures and other media included with this pamphlet may be under different copyright restrictions.

Design by: Mihai Toma, Nick White and Sean Worley

Printed by: Rope Press, Birmingham

This pamphlet is published in a series of 7 as part of the Radical Open Access II – The Ethics of Care conference, which took place June 26-27 at Coventry University. More information about this conference and about the contributors to this pamphlet can be found at:
http://radicaloa.co.uk/conferences/ROA2

This pamphlet was made possible due to generous funding from the arts and humanities research studio, The Post Office, a project of Coventry University's Centre for Postdigital Cultures and due to the combined efforts of authors, editors, designers and printers.

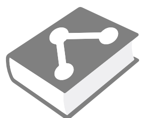
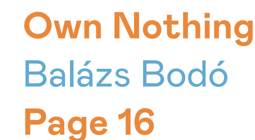


Table of Contents

Guerrilla Open Access: Terms Of Struggle
Memory of the World
Page 4

Recursive Publics and Open Access
Christopher Kelty
Page 6



Own Nothing
Balázs Bodó
Page 16

What if We Aren't the Only Guerrillas Out There?
Laurie Allen
Page 26

Guerilla Open Access: Terms Of Struggle

In the 1990s, the Internet offered a horizon from which to imagine what society could become, promising autonomy and self-organization next to redistribution of wealth and collectivized means of production. While the former was in line with the dominant ideology of freedom, the latter ran contrary to the expanding enclosures in capitalist globalization. This antagonism has led to epochal copyfights, where free software and piracy kept the promise of radical commoning alive.

Free software, as Christopher Kelty writes in this pamphlet, provided a model 'of a shared, collective, process of making software, hardware and infrastructures that cannot be appropriated by others'. Well into the 2000s, it served as an inspiration for global free culture and open access movements who were speculating that distributed infrastructures of knowledge production could be built, as the Internet was, on top of free software.

For a moment, the hybrid world of ad-financed Internet giants—sharing code, advocating open standards and interoperability—and users empowered by these services, convinced almost everyone that a new reading/writing culture was possible. Not long after the crash of 2008, these disruptors, now wary monopolists, began to ingest smaller disruptors and close off their platforms. There was still free software somewhere underneath, but without the 'original sense of shared, collective, process'. So, as Kelty suggests, it was hard to imagine that for-profit academic publishers wouldn't try the same with open access.

Heeding Aaron Swartz's call to civil disobedience, Guerrilla Open Access has emerged out of the outrage over digitally-enabled enclosure of knowledge that has allowed these for-profit academic publishers to appropriate extreme profits that stand in stark contrast to the cuts, precarity, student debt and asymmetries of access in education. Shadow libraries stood in for the access denied to public libraries, drastically reducing global asymmetries in the process.

This radicalization of access has changed how publications travel across time and space. Digital archiving, cataloging and sharing is transforming what we once considered as private libraries. Amateur librarianship is becoming public shadow librarianship. Hybrid use, as poetically unpacked in Balázs Bodó's reflection on his own personal library, is now entangling print and digital in novel ways. And, as he warns, the terrain of antagonism is shifting. While for-profit publishers are seemingly conceding to Guerrilla Open Access, they are opening new territories: platforms centralizing data, metrics and workflows, subsuming academic autonomy into new processes of value extraction.

The 2010s brought us hope and then realization how little digital networks could help revolutionary movements. The redistribution toward the wealthy, assisted by digitization, has eroded institutions of solidarity. The embrace of privilege—marked by misogyny, racism and xenophobia—this has catalyzed is nowhere more evident than in the climate denialism of the Trump administration. Guerrilla archiving of US government climate change datasets, as recounted by Laurie Allen, indicates that more technological innovation simply won't do away with the 'post-truth' and that our institutions might be in need of revision, replacement and repair.

As the contributions to this pamphlet indicate, the terms of struggle have shifted: not only do we have to continue defending our shadow libraries, but we need to take back the autonomy of knowledge production and rebuild institutional grounds of solidarity.

Memory of the World
http://memoryoftheworld.org

Recursive Publics and Open Access

Christopher Kelty

Ten years ago, I published a book called *Two Bits: The Cultural Significance of Free Software* (Kelty 2008).[1] Duke University Press and my editor Ken Wissoker were enthusiastically accommodating of my demands to make the book freely and openly available. They also played along with my desire to release the 'source code' of the book (i.e. HTML files of the chapters), and to compare the data on readers of the open version to print customers. It was a moment of exploration for both scholarly presses and for me. At the time, few authors were doing this other than Yochai Benkler (2007) and Cory Doctorow[2], both activists and advocates for free software and open access (OA), much as I have been. We all shared, I think, a certain fanaticism of the convert that came from recognizing free software as an historically new, and radically different mode of organizing economic and political activity. *Two Bits* gave me a way to talk not only about free software, but about OA and the politics of the university (Kelty et al. 2008; Kelty 2014). Ten years later, I admit to a certain pessimism at the way things have turned out. The promise of free software has foundered, though not disappeared, and the question of what it means to achieve the goals of OA has been swamped by concerns about costs, arcane details of repositories and versioning, and ritual offerings to the metrics God.

When I wrote *Two Bits*, it was obvious to me that the collectives who built free software were *essential* to the very structure and operation of a standardized Internet. Today, free software and 'open source' refer to dramatically different constellations of practice and people. Free software gathers around itself those committed to the original sense of a shared, collective, process of making software, hardware and infrastructures that cannot be appropriated by others. In political terms, I have always identified free software with a very specific, updated, version of classical Millian liberalism. It sustains a belief in the capacity for collective action and rational thought as aids to establishing a flourishing human livelihood. Yet it also preserves an outdated blind faith in the automatic functioning of meritorious speech, that the best ideas will inevitably rise to the top. It is an updated classical liberalism that saw in software and networks a new place to resist the tyranny of the conventional and the taken for granted.

By contrast, open source has come to mean something quite different: an ecosystem controlled by an oligopoly of firms which maintains a shared pool of components and frameworks that lower the costs of education, training, and software creation in the service of establishing winner-take-all platforms. These are built on open source, but they do not carry the principles of freedom or openness all the way through to the platforms themselves.[3] What open source has become is now almost the opposite of free software—it is authoritarian, plutocratic, and nepotistic, everything liberalism wanted to resist. For example, precarious labor and platforms such as Uber or Task Rabbit are built upon and rely on the fruits of the labor of 'open source', but the platforms that result do not follow the same principles—they are not open or free in any meaningful sense—to say nothing of the Uber drivers or task rabbits who live by the platforms.

Does OA face the same problem? In part, my desire to 'free the source' of my book grew out of the unfinished business of digitizing the scholarly record. It is an irony that much of the work that went into designing the Internet at its outset in the 1980s, such as gopher, WAIS, and the HTML of CERN, was conducted in the name of the digital transformation of the library. But by 2007, these aims were swamped by attempts to transform the Internet into a giant factory of data extraction. Even in 2006-7 it was clear that this unfinished business of digitizing the scholarly record was going to become a problem—both because it was being overshadowed by other concerns, and because of the danger it would eventually be subjected to the very platformization underway in other realms.

Because if the platform capitalism of today has ended up being parasitic on the free software that enabled it, then why would this not also be true of scholarship more generally? Are we not witnessing a transition to a world where scholarship is directed—in its very content and organization—towards the profitability of the platforms that ostensibly serve it?[4] Is it not possible that the platforms created to 'serve science'—Elsevier's increasing acquisition of tools to control the entire life-cycle of research, or ResearchGate's ambition to become the single source for all academics to network and share research—that these platforms might actually end up warping the very content of scholarly production in the service of their profitability?

To put this even more clearly: OA has come to exist and scholarship is more available and more widely distributed than ever before. But, scholars now have less control, and have taken less responsibility for the means of production of scientific research, its circulation, and perhaps even the content of that science.

The Method of Modulation

When I wrote *Two Bits* I organized the argument around the idea of *modulation*: free software is simply one assemblage of technologies, practices, and people aimed at resolving certain problems regarding the relationship between knowledge (or software tools related to knowledge) and power (Hacking 2004; Rabinow 2003). Free software as such was and still is changing as each of its elements evolve or are recombined. Because OA derives some of its practices directly from free software, it is possible to observe how these different elements have been worked over in the recent past, as well as how new and surprising elements are combined with OA to transform it. Looking back on the elements I identified as central to free software, one can ask: how is OA different, and what new elements are modulating it into something possibly unrecognizable?

Sharing source code

Shareable source code was a concrete and necessary achievement for free software to be possible. Similarly, the necessary ability to circulate digital texts is a significant achievement—but such texts are shareable in a much different way. For source code, computable streams of text are everything—anything else is a 'blob' like an image, a video or any binary file. But scholarly texts are blobs: Word or Portable Document Format (PDF) files. What's more, while software programmers may love 'source code', academics generally hate it—anything less than the final, typeset version is considered unfinished (see e.g. the endless disputes over 'author's final versions' plaguing OA).[5] Finality is important. Modifiability of a text, especially in the humanities and social sciences, is acceptable only when it is an experiment of some kind.

In a sense, the source code of science is not a code at all, but a more abstract set of relations between concepts, theories, tools, methods, and the disciplines and networks of people who operate with them, critique them, extend them and try to maintain control over them even as they are shared within these communities.

Defining openness

For free software to make sense as a solution, those involved first had to characterize the problem it solved—and they did so by identifying a pathology in the worlds of corporate capitalism and engineering in the 1980s: that computer corporations were closed organizations who re-invented basic tools and infrastructures in a race to dominate a market. An 'open system,' by contrast, would

avoid the waste of 'reinventing the wheel' and of pathological competition, allowing instead modular, reusable parts that could be modified and recombined to build better things in an upward spiral of innovation. The 1980s ideas of modularity, modifiability, abstraction barriers, interchangeable units have been essential to the creation of digital infrastructures.

To propose an 'open science' thus modulates this definition—and the idea works in some sciences better than others. Aside from the obviously different commercial contexts, philosophers and literary theorists just don't think about openness this way—theories and arguments may be used as building blocks, but they are not modular in quite the same way. Only the free circulation of the work, whether for recombination or for reference and critique, remains a *sine qua non* of the theory of openness proposed there. It is opposed to a system where it is explicit that only certain people have access to the texts (whether that be through limitations of secrecy, or limitations on intellectual property, or an implicit elitism).

Writing and using copyright licenses

Of all the components of free software that I analyzed, this is the one practice that remains the least transformed—OA texts use the same CC licenses pioneered in 2001, which were a direct descendant of free software licenses.

A novel modulation of these licenses is the OA policies (the embrace of OA in Brazil for instance, or the spread of OA Policies starting with Harvard and the University of California, and extending to the EU Mandate from 2008 forward). Today the ability to control the circulation of a text with IP rights is far less economically central to the strategies of publishers than it was in 2007, even if they persist in attempting to do so. At the same time, funders, states, and universities have all adopted patchwork policies intended to both sustain green OA, and push publishers to innovate their own business models in gold and hybrid OA. While green OA is a significant success on paper, the actual use of it to circulate work pales

in comparison to the commercial control of circulation on the one hand, and the increasing success of shadow libraries on the other. Repositories have sprung up in every shape and form, but they remain largely *ad hoc*, poorly coordinated, and underfunded solutions to the problem of OA.

Coordinating collaborations

The *collective* activity of free software is ultimately the most significant of its achievements—marrying a form of intensive small-scale interaction amongst programmers, with sophisticated software for managing complex objects (version control and GitHub-like sites). There has been constant innovation in these tools for controlling, measuring, testing, and maintaining software.

By contrast, the collective activity of scholarship is still largely a pre-modern affair. It is coordinated largely by the idea of 'writing an article together' and not by working to maintain some larger map of what a research topic, community, or discipline has explored—what has worked and what has not.

This focus on the coordination of collaboration seemed to me to be one of the key advantages of free software, but it has turned out to be almost totally absent from the practice or discussion of OA. Collaboration and the recombination of elements of scholarly practice obviously happens, but it does not depend on OA in any systematic way: there is only the counterfactual that without it, many different kinds of people are excluded from collaboration or even simple participation in, scholarship, something that most active scholars are willfully ignorant of.

Fomenting a movement

I demoted the idea of a social movement to merely one component of the success of free software, rather than let it be—as most social scientists would have it—the principal container for free software. They are not the whole story.

Is there an OA movement? Yes and no. Librarians remain the most activist and organized. The handful of academics who care about it have shifted to caring about it in primarily a bureaucratic sense, forsaking the cross-organizational aspects of a movement in favor of activism within universities (to which I plead guilty). But this transformation forsakes the need for addressing the collective, collaborative responsibility for scholarship in favor of letting individual academics, departments, and disciplines be the focus for such debates.

By contrast, the publishing industry works with a phantasmatic idea of both an OA 'movement' and of the actual practices of scholarship—they too defer, in speech if not in practice, to the academics themselves, but at the same time must create tools, innovate processes, establish procedures, acquire tools and companies and so on in an effort to capture these phantasms and to prevent academics from collectively doing so on their own.

And what new components? The five above were central to free software, but OA has other components that are arguably more important to its organization and transformation.

Money, i.e. library budgets

Central to almost all of the politics and debates about OA is the political economy of publication. From the 'bundles' debates of the 1990s to the gold/green debates of the 2010s, the sole source of money for publication long ago shifted into the library budget. The relationship that library budgets have to other parts of the political economy of research (funding for research itself, debates about tenured/non-tenured, adjunct and other temporary salary structures) has shifted as a result of the demand for OA, leading libraries to re-conceptualize themselves as potential publishers, and publishers to re-conceptualize themselves as serving 'life cycles' or 'pipeline' of research, not just its dissemination.

Metrics

More than anything, OA is promoted as a way to continue to feed the metrics God. OA means more citations, more easily computable data, and more visible uses and re-uses of publications (as well as 'open data' itself, when conceived of as product and not measure). The innovations in the world of metrics—from the quiet expansion of the platforms of the publishers, to the invention of 'alt metrics', to the enthusiasm of 'open science' for metrics-driven scientific methods—forms a core feature of what 'OA' is today, in a way that was not true of free software before it, where metrics concerning users, downloads, commits, or lines of code were always after-the-fact measures of quality, and not constitutive ones.

Other components of this sort might be proposed, but the main point is to resist to clutch OA as if it were the beating heart of a social transformation in science, as if it were a *thing* that must exist, rather than a configuration of elements at a moment in time. OA was a solution—but it is too easy to lose sight of the problem.

Open Access without Recursive Publics

When we no longer have any commons, but only platforms, will we still have knowledge as we know it? This is a question at the heart of research in the philosophy and sociology of knowledge—not just a concern for activism or social movements. If knowledge is socially produced and maintained, then the nature of the social bond surely matters to the nature of that knowledge. This is not so different than asking whether we will still have labor or work, as we have long known it, in an age of precarity? What is the knowledge equivalent of precarity (i.e. not just the existence of precarious knowledge workers, but a kind of *precarious knowledge* as such)?

Do we not already see the evidence of this in the 'post-truth' of fake news, or the deliberate refusal by those in power to countenance evidence, truth, or established systems of argument and debate? The relationship between

knowledge and power is shifting dramatically, because the costs—and the stakes—of producing high quality, authoritative knowledge have also shifted. It is not so powerful any longer; science does not speak truth to power because truth is no longer so obviously important to power.

Although this is a pessimistic portrait, it may also be a sign of something yet to come. Free software as a community, has been and still sometimes is critiqued as being an exclusionary space of white male sociality (Nafus 2012; Massanari 2016; Ford and Wajcman 2017; Reagle 2013). I think this critique is true, but it is less a problem of identity than it is a pathology of a certain form of liberalism: a form that demands that merit consists only in the *content* of the things we say (whether in a political argument, a scientific paper, or a piece of code), and not in the ways we say them, or who is encouraged to say them and who is encouraged to remain silent (Dunbar-Hester 2014).

One might, as a result, choose to throw out liberalism altogether as a broken philosophy of governance and liberation. But it might also be an opportunity to focus much more specifically on a particular problem of liberalism, one that the discourse of OA also relies on to a large extent. Perhaps it is not the case that merit derives solely from the content of utterances freely and openly circulated, but also from the *ways in which they are uttered, and the dignity of the people who utter them*. An OA (or a free software) that embraced that principle would demand that we pay attention to different problems: how are our platforms, infrastructures, tools organized and built to support not just the circulation of putatively true statements, but the ability to say them in situated and particular ways, with respect for the dignity of who is saying them, and with the freedom to explore the limits of *that* kind of liberalism, should we be so lucky to achieve it.

[1] https://twobits.net/download/index.html

[2] https://craphound.com/

[3] For example, Platform Cooperativism https://platform.coop/directory

[4] See for example the figure from 'Rent Seeking by Elsevier,' by Alejandro Posada and George Chen (http://knowledgegap.org/index.php/sub-projects/rent-seeking-and-financialization-of-the-academic-publishing-industr preliminary-findings/)

[5] See Sherpa/Romeo http://www.sherpa.ac.uk/romeo/index.php

References

Benkler, Yochai. 2007. *The Wealth of Networks: How Social Production Transforms Markets and Freedom.* Yale University Press.

Dunbar-Hester, Christina. 2014. *Low Power to the People: Pirates, Protest, and Politics in FM Radio Activism.* MIT Press.

Ford, Heather, and Judy Wajcman. 2017. "'Anyone Can Edit', Not Everyone Does: Wikipedia's Infrastructure and the Gender Gap". *Social Studies of Science* 47 (4): 511–527. doi:10.1177/0306312717692172.

Hacking, I. 2004. *Historical Ontology.* Harvard University Press.

Kelty, Christopher M. 2014. "Beyond Copyright and Technology: What Open Access Can Tell Us About Precarity, Authority, Innovation, and Automation in the University Today". *Cultural Anthropology* 29 (2): 203–215. doi:10.14506/ca29.2.02.

———. 2008. *Two Bits: The Cultural Significance of Free Software.* Durham, N.C.: Duke University Press.

Kelty, Christopher M., et al. 2008. "Anthropology In/of Circulation: a Discussion". *Cultural Anthropology* 23 (3).

Massanari, Adrienne. 2016. "#gamergate and the Fappening: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures". *New Media & Society* 19 (3): 329–346. doi:10.1177/1461444815608807.

Nafus, Dawn. 2012. "'Patches don't have gender': What is not open in open source software". *New Media & Society* 14, no. 4: 669–683. Visited on 04/01/2014. http://doi:10.1177/1461444811422887.

Rabinow, Paul. 2003. *Anthropos Today: Reflections on Modern Equipment.* Princeton University Press.

Reagle, Joseph. 2013. ""Free As in Sexist?" Free Culture and the Gender Gap". *First Monday* 18 (1). doi:10.5210/fm.v18i1.4291.

Own Nothing

Balázs Bodó

Flow My Tears

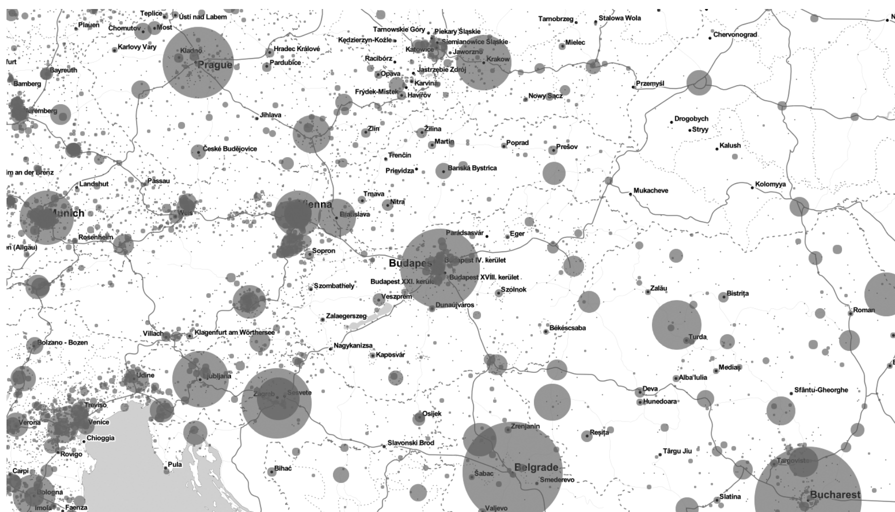
My tears cut deep grooves into the dust on my face. Drip, drip, drop, they hit the floor and disappear among the torn pages scattered on the floor.

This year it dawned on us that we cannot postpone it any longer: our personal library has to go. Our family moved countries more than half a decade ago, we switched cultures, languages, and chose another future. But the past, in the form of a few thousand books in our personal library, was still neatly stacked in our old apartment, patiently waiting, books that we bought and enjoyed — and forgot; books that we bought and never opened; books that we inherited from long-dead parents and half-forgotten friends. Some of them were important. Others were relevant at one point but no longer, yet they still reminded us who we once were.

When we moved, we took no more than two suitcases of personal belongings. The books were left behind. The library was like a sick child or an ailing parent, it hung over our heads like an unspoken threat, a curse. It was clear that sooner or later something had to be done about it, but none of the options available offered any consolation. It made no sense to move three thousand books to the other side of this continent. We decided to emigrate, and not to take our past with us, abandon the contexts we were fleeing from. We made a choice to leave behind the history, the discourses, the problems and the pain that accumulated in the books of our library. I knew exactly what it was I didn't want to teach to my children once we moved. So we did not move the books. We pretended that we would never have to think about what this decision really meant. Up until today. This year we needed to empty the study with the shelves. So I'm standing in our library now, the dust covering my face, my hands, my clothes. In the middle of the floor there are three big crates and one small box. The small box swallows what we'll ultimately take with us, the books I want to show to my son when he gets older, in case he still wants to read. One of the big crates will be taken away by the antiquarian. The other will be given to the school library next door. The third is the wastebasket, where everything else will ultimately go.

Drip, drip, drip, my tears flow as I throw the books into this last crate, drip, drip, drop. Sometimes I look at my partner, working next to me, and I can see on her face that she is going through the same emotions. I sometimes catch the sight of her trembling hand, hesitating for a split second where a book should ultimately go, whether we could, whether we should save that particular one, because… But we either save them all or we are as ruthless as all those millions of people throughout history, who had an hour to pack their two suitcases before they needed to leave. Do we truly need this book? Is this a book we'll want to read? Is this book an inseparable part of our identity? Did we miss this book at all in the last five years? Is this a text I want to preserve for the future, for potential grandchildren who may not speak my mother tongue at all? What is the function of the book? What is the function of this particular book in my life? Why am I hesitating throwing it out? Why should I hesitate at all? Drop, drop, drop, a decision has been made. Drop, drop, drop, books are falling to the bottom of the crates.

We are killers, gutting our library. We are like the half-drown sailor, who got entangled in the ropes, and went down with the ship, and who now frantically tries to cut himself free from the detritus that prevents him to reach the freedom of the surface, the sunlight and the air.



Own Nothing, Have Everything

Do you remember Napster's slogan after it went legit, trying to transform itself into a legal music service around 2005? 'Own nothing, have everything' – that was the headline that was supposed to sell legal streaming music. How stupid, I thought. How could you possibly think that lack of ownership would be a good selling point? What does it even mean to 'have everything' without ownership? And why on earth would not everyone want to own the most important constituents of their own self, their own identity? The things I read, the things I sing, make me who I am. Why wouldn't I want to own these things?

How revolutionary this idea had been I reflected as I watched the local homeless folks filling up their sacks with the remains of my library. How happy I would be if I could have all this stuff I had just thrown away without actually having to own any of it. The proliferation of digital texts led me to believe that we won't be needing dead wood libraries at all, at least no more than we need vinyl to listen to, or collect music. There might be geeks, collectors, specialists, who for one reason or another still prefer the physical form to the digital, but for the rest of us convenience, price, searchability, and all the other digital goodies give enough reason not to collect stuff that collects dust.

I was wrong to think that. I now realize that the future is not fully digital, it is more a physical-digital hybrid, in which the printed book is not simply an endangered species protected by a few devoted eccentrics who refuse to embrace the obvious

advantages of a fully digital book future. What I see now is the emergence of a strange and shapeshifting-hybrid of diverse physical and electronic objects and practices, where the relative strengths and weaknesses of these different formats nicely complement each other.

This dawned on me after we had moved into an apartment without a bookshelf. I grew up in a flat that housed my parents' extensive book collection. I knew the books by their cover and from time to time something made me want to take it from the shelf, open it and read it. This is how I discovered many of my favorite books and writers. With the e-reader, and some of the best shadow libraries at hand, I felt the same at first. I felt liberated. I could experiment without cost or risk, I could start—or stop—a book, I didn't have to consider the cost of buying and storing a book that was ultimately not meant for me. I could enjoy the books without having to carry the burden and responsibility of ownership.

Did you notice how deleting an epub file gives you a different feeling than throwing out a book? You don't have to feel guilty, you don't have to feel anything at all.
So I was reading, reading, reading like never before. But at that time my son was too young to read, so I didn't have to think about him, or anyone else besides myself. But as he was growing, it slowly dawned on me: without these physical books how will I be able to give him the same chance of serendipity, and of discovery, enchantment, and immersion that I got in my father's library? And even later, what will I give him as his heritage? Son, look into this folder of PDFs: this is my legacy, your heritage, explore, enjoy, take pride in it?

Collections of anything, whether they are art, books, objects, people, are inseparable from the person who assembled that collection, and when that person is gone, the collection dies, as does the most important inroad to it: the will that created this particular order of things has passed away. But the heavy and unavoidable physicality of a book collection forces all those left behind to make an effort to approach, to force their way into, and try to navigate that garden of forking paths that is *someone else's library*. Even if you ultimately get rid of everything, you have to introduce yourself to every book, and let every book introduce itself to you, so you know what you're throwing out. Even if you'll ultimately kill, you will need to look into the eyes of all your victims.

With a digital collection that's, of course, not the case.

The e-book is ephemeral. It has little past and even less chance to preserve the fingerprints of its owners over time. It is impersonal, efficient, fast, abundant, like

fast food or plastic, it flows through the hand like sand. It lacks the embodiment, the materiality which would give it a life in a temporal dimension. If you want to network the dead and the unborn, as is the ambition of every book, then you need to print and bind, and create heavy objects that are expensive, inefficient and a burden. This burden subsiding in the object is the bridge that creates the intergenerational dimension, that forces you to think of the value of a book.

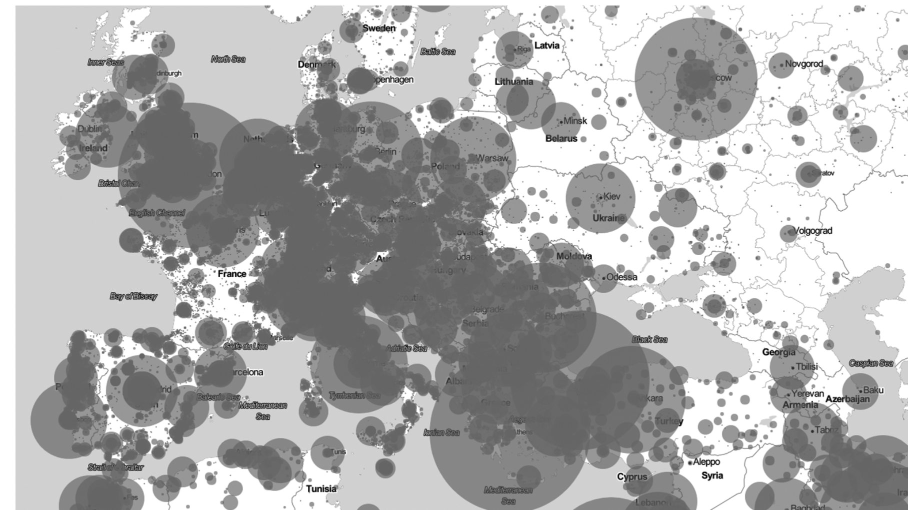
Own nothing, have nothing. Own everything, and your children will hate you when you die.

I have to say, I'm struggling to find a new balance here. I started to buy books again, usually books that I'd already read from a stolen copy on-screen. I know what I want to buy, I know what is worth preserving. I know what I want to show to my son, what I want to pass on, what I would like to take care of over time. Before, book buying for me was an investment into a stranger. Now that thrill is gone forever. I measure up the merchandise well beforehand, I build an intimate relationship, we make love again and again, before moving in together.

It is certainly a new kind of relationship with the books I bought since I got my e-reader. I still have to come to terms with the fact that the books I bought this way are rarely opened, as I already know them, and their role is not to be read, but to be together. What do I buy, and what do I get? Temporal, existential security? The chance of serendipity, if not for me, then for the people around me? The reassuring materiality of the intimacy I built with these texts through another medium?

All of these and maybe more. But in any case, I sense that this library, the physical embodiment of a physical-electronic hybrid collection with its unopened books and overflowing e-reader memory cards, is very different from the library I had, and the library I'm getting rid of at this very moment. The library that I inherited, the library that grew organically from the detritus of the everyday, the library that accumulated books similar to how the books accumulated dust, as is the natural way of things, this library was full of unknowns, it was a library of potentiality, of opportunities, of trips waiting to happen. This new, hybrid library is a collection of things that I'm familiar with. I intimately know every piece, they hold little surprise, they offer few discoveries — at least for me. The exploration, the discovery, the serendipity, the pre-screening takes place on the e-reader, among the ephemeral, disposable PDFs and epubs.

Have everything, and own a few.



We Won

This new hybrid model is based on the cheap availability of digital books. In my case, the free availability of pirated copies available through shadow libraries. These libraries don't have everything on offer, but they have books in an order of magnitude larger than I'll ever have the time and chance to read, so they offer enough, enough for me to fill up hard drives with books I want to read, or at least skim, to try, to taste. As if I moved into an infinite bookstore or library, where I can be as promiscuous, explorative, nomadic as I always wanted to be. I can flirt with books, I can have a quickie, or I can leave them behind without shedding a single tear.

I don't know how this hybrid library, and this analogue-digital hybrid practice of reading and collecting would work without the shadow libraries which make everything freely accessible. I rely on their supply to test texts, and feed and grow my print library. E-books are cheaper than their print versions, but they still cost money, carry a risk, a cost of experimentation. Book-streaming, the flat-rate, the all-you-can-eat format of accessing books is at the moment only available to audiobooks, but rarely for e-books. I wonder why.

Did you notice that there are no major book piracy lawsuits?

Of course there is the lawsuit against Sci-Hub and Library Genesis in New York, and there is another one in Canada against aaaaarg, causing major nuisance to those who have been named in these cases. But this is almost negligible compared to the high profile wars the music and audiovisual industries waged against Napster, Grokster, Kazaa, megaupload and their likes. It is as if book publishers have completely given up on trying to fight piracy in the courts, and have launched a few lawsuits only to maintain the appearance that they still care about their digital copyrights. I wonder why.

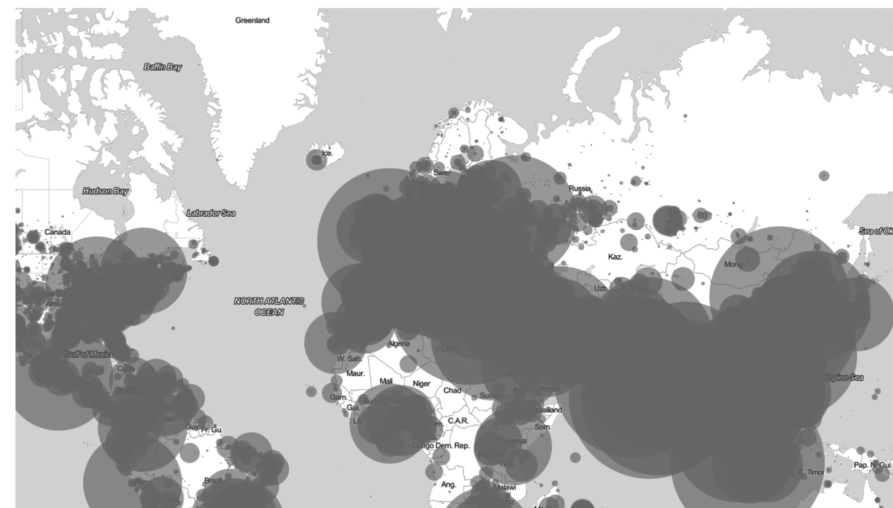
I know the academic publishing industry slightly better than the mainstream popular fiction market, and I have the feeling that in the former copyright-based business models are slowly being replaced by something else. We see no major anti-piracy efforts from publishers, not because piracy is non-existent — on the contrary, it is global, and it is big — but because the publishers most probably realized that in the long run the copyright-based exclusivity model is unsustainable. The copyright wars of the last two decades taught them that law cannot put an end to piracy. As the Sci-Hub case demonstrates, you can win all you want in a New York court, but this has little real-world effect as long as the conditions that attract the users to the shadow libraries remain.

Exclusivity-based publishing business models are under assault from other sides as well. Mandated open access in the US and in the EU means that there is a quickly growing body of new research for the access of which publishers cannot charge money anymore. LibGen and Sci-Hub make it harder to charge for the back catalogue. Their sheer existence teaches millions on what uncurtailed open access really is, and makes it easier for university libraries to negotiate with publishers, as they don't have to worry about their patrons being left without any access at all.

The good news is that radical open access may well be happening. It is a less and less radical idea to have things freely accessible. One has to be less and less radical to achieve the openness that has been long overdue. Maybe it is not yet obvious today and the victory is not yet universal, maybe it'll take some extra years, maybe it won't ever be evenly distributed, but it is obvious that this genie, these millions of books on everything from malaria treatments to critical theory, cannot be erased, and open access will not be undone, and the future will be free of access barriers.

Who is downloading books and articles? Everyone. Radical open access? We won, if you like.

Drip, drip, drop, its only nostalgia. My heart is light, as I don't have to worry about gutting the library. Soon it won't matter at all.



We Are Not Winning at All

But did we really win? If publishers are happy to let go of access control and copyright, it means that they've found something that is even more profitable than selling back to us academics the content that we have produced. And this more profitable something is of course data. Did you notice where all the investment in academic publishing went in the last decade? Did you notice SSRN, Mendeley, Academia.edu, ScienceDirect, research platforms, citation software, manuscript repositories, library systems being bought up by the academic publishing industry? All these platforms and technologies operate on and support open access content, while they generate data on the creation, distribution, and use of knowledge; on individuals, researchers, students, and faculty; on institutions, departments, and programs. They produce data on the performance, on the success and the failure of the whole domain of research and education. This is the data that is being privatized, enclosed, packaged, and sold back to us.

Taylorism reached academia. In the name of efficiency, austerity, and transparency, our daily activities are measured, profiled, packaged, and sold to the highest bidder. But in this process of quantification, knowledge on ourselves is lost for us, unless we pay. We still have some patchy datasets on what we do, on who we are, we still have this blurred reflection in the data-mirrors that we still do control. But this path of self-enlightenment is quickly waning as less and less data sources about us are freely available to us.

I strongly believe that information on the self is the foundation of self-determination. We need to have data on how we operate, on what we do in order to know who we are. This is what is being privatized away from the academic community, this is being taken away from us.

Radical open access. Not of content, but of the data about ourselves. This is the next challenge. We will digitize every page, by hand if we must, that process cannot be stopped anymore. No outside power can stop it and take that from us. Drip, drip, drop, this is what I console myself with, as another handful of books land among the waste.

But the data we lose now will not be so easy to reclaim.

What if We Aren't the Only Guerrillas Out There?

Laurie Allen

My goal in this paper is to tell the story of a grass-roots project called Data Refuge (http://www.datarefuge.org) that I helped to co-found shortly after, and in response to, the Trump election in the USA. Trump's reputation as anti-science, and the promise that his administration would elevate people into positions of power with a track record of distorting, hiding, or obscuring the scientific evidence of climate change caused widespread concern that valuable federal data was now in danger. The Data Refuge project grew from the work of Professor Bethany Wiggin and the graduate students within the Penn Program in Environmental Humanities (PPEH), notably Patricia Kim, and was formed in collaboration with the Penn Libraries, where I work. In this paper, I will discuss the Data Refuge project, and call attention to a few of the challenges inherent in the effort, especially as they overlap with the goals of this collective. I am not a scholar. Instead, I am a librarian, and my perspective as a practicing informational professional informs the way I approach this paper, which weaves together the practical and technical work of 'saving data' with the theoretical, systemic, and ethical issues that frame and inform what we have done.

I work as the head of a relatively small and new department within the libraries of the University of Pennsylvania, in the city of Philadelphia, Pennsylvania, in the US. I was hired to lead the Digital Scholarship department in the spring of 2016, and most of the seven (soon to be eight) people within Digital Scholarship joined the library since then in newly created positions. Our group includes a mapping and spatial data librarian and three people focused explicitly on supporting the creation of new Digital Humanities scholarship. There are also two people in the department who provide services connected with digital scholarly open access publishing, including the maintenance of the Penn Libraries' repository of open access scholarship, and one Data Curation and Management Librarian. This Data Librarian, Margaret Janz, started working with us in September 2016, and features heavily into the story I'm about to tell about our work helping to build Data Refuge. While Margaret and I were the main people in our department involved in the project, it is useful to understand the work we did as connected more broadly to the intersection of activities—from multimodal, digital, humanities creation to open access publishing across disciplines—represented in our department in Penn.

At the start of Data Refuge, Professor Wiggin and her students had already been exploring the ways that data about the environment can empower communities through their art, activism, and research, especially along the lower Schuylkill River in Philadelphia. They were especially attuned to the ways that missing data, or data that is not collected or communicated, can be a source of disempowerment. After the Trump election, PPEH graduate students raised the concern that the political commitments of the new administration would result in the disappearance of environmental and climate data that is vital to work in cities and communities around the world. When they raised this concern with the library, together we co-founded Data Refuge. It is notable to point out that, while the Penn Libraries is a large and relatively well-resourced research library in the United States, it did not have any automatic way to ingest and steward the data that Professor Wiggin and her students were concerned about. Our system of acquiring, storing, describing and sharing publications did not account for, and could not easily handle, the evident need to take in large quantities of public data from the open web and make them available and citable by future scholars. Indeed, no large research library was positioned to respond to this problem in a systematic way, though there was general agreement that the community would like to help.

The collaborative, grass-roots movement that formed Data Refuge included many librarians, archivists, and information professionals, but it was clear from the beginning that my own profession did not have in place a system for stewarding these vital information resources, or for treating them as 'publications' of the

federal government. This fact was widely understood by various members of our profession, notably by government document librarians, who had been calling attention to this lack of infrastructure for years. As Government Information Librarian Shari Laster described in a blog post in November of 2016, government documents librarians have often felt like they are 'under siege' not from political forces, but from the inattention to government documents afforded by our systems and infrastructure. Describing the challenges facing the profession in light of the 2016 election, she commented: "Government documents collections in print are being discarded, while few institutions are putting strategies in place for collecting government information in digital formats. These strategies are not expanding in tandem with the explosive proliferation of these sources, and certainly not in pace with the changing demands for access from public users, researchers, students, and more." (Laster 2016) Beyond government documents librarians, our project joined efforts that were ongoing in a huge range of communities, including: open data and open science activists; archival experts working on methods of preserving born-digital content; cultural historians; federal data producers and the archivists and data scientists they work with; and, of course, scientists.

Born from the collaboration between Environmental Humanists and Librarians, Data Refuge was always an effort both at storytelling and at storing data. During the first six months of 2017, volunteers across the US (and elsewhere) organized more than 50 Data Rescue events, with participants numbering in the thousands. At each event, a group of volunteers used tools created by our collaborators at the Environmental and Data Governance Initiative (EDGI) (https://envirodatagov.org/) to support the End of Term Harvest (http://eotarchive.cdlib.org/) project by identifying seeds from federal websites for web archiving in the Internet Archive. Simultaneously, more technically advanced volunteers wrote scripts to pull data out of complex data systems, and packaged that data for longer term storage in a repository we maintained at datarefuge.org. Still other volunteers held teach-ins, built profiles of data storytellers, and otherwise engaged in safeguarding environmental and climate data through community action (see http://www.ppehlab.org/datarefugepaths). The repository at datarefuge.org that houses the more difficult data sources has been stewarded by myself and Margaret Janz through our work at Penn Libraries, but it exists outside the library's main technical infrastructure.[1]

This distributed approach to the work of downloading and saving the data encouraged people to see how they were invested in environmental and scientific data, and to consider how our government records should be considered the property of all of us. Attending Data Rescue events was a way for people who value

the scientific record to fight back, in a concrete way, against an anti-fact establishment. By downloading data and moving it into the Internet Archive and the Data Refuge repository, volunteers were actively claiming the importance of accurate records in maintaining or creating a just society.

Of course, access to data need not rely on its inclusion in a particular repository. As is demonstrated so well in other contexts, technological methods of sharing files can make the digital repositories of libraries and archives seem like a redundant holdover from the past. However, as I will argue further in this paper, the data that was at risk in Data Refuge differed in important ways from the contents of what Bodó refers to as 'shadow libraries' (Bodó 2015). For opening access to copies of journals articles, shadow libraries work perfectly. However, the value of these shadow libraries relies on the existence of the widely agreed upon trusted versions. If in doubt about whether a copy is trustworthy, scholars can turn to more mainstream copies, if necessary. This was not the situation we faced building Data Refuge. Instead, we were often dealing with the sole public, authoritative copy of a federal dataset and had to assume that, if it were taken down, there would be no way to check the authenticity of other copies. The data was not easily pulled out of systems as the data and the software that contained them were often inextricably linked. We were dealing with unique, tremendously valuable, but often difficult-to-untangle datasets rather than neatly packaged publications. The workflow we established was designed to privilege authenticity and trustworthiness over either the speed of the copying or the easy usability of the resulting data.[2] This extra care around authenticity was necessary because of the politicized nature of environmental data that made many people so worried about its removal after the election. It was important that our project supported the strongest possible scientific arguments that could be made with the data we were 'saving'. That meant that our copies of the data needed to be citable in scientific scholarly papers, and that those citations needed to be able to withstand hostile political forces who claim that the science of human-caused climate change is 'uncertain'. It

was easy to imagine in the Autumn of 2016, and even easier to imagine now, that hostile actors might wish to muddy the science of climate change by releasing fake data designed to cast doubt on the science of climate change. For that reasons, I believe that the unique facts we were seeking to safeguard in the Data Refuge bear less similarity to the contents of shadow libraries than they do to news reports in our current distributed and destabilized mass media environment. Referring to the ease of publishing ideas on the open web, Zeynep Tufecki wrote in a recent column, "And sure, it is a golden age of free speech—if you can believe your lying eyes. Is that footage you're watching real? Was it really filmed where and when it says it was? Is it being shared by alt-right trolls or a swarm of Russian bots? Was it maybe even generated with the help of artificial intelligence? (Yes, there are systems that can create increasingly convincing fake videos.)" (Tufekci 2018). This was the state we were trying to avoid when it comes to scientific data, fearing that we might have the only copy of a given dataset without solid proof that our copy matched the original.

If US federal websites cease functioning as reliable stewards of trustworthy scientific data, reproducing their data without a new model of quality control risks producing the very censorship that our efforts are supposed to avoid, and further undermining faith in science. Said another way, if volunteers duplicated federal data all over the Internet without a trusted system for ensuring the authenticity of that data, then as soon as the originals were removed, a sea of fake copies could easily render the original invisible, and they would be just as effectively censored. "The most effective forms of censorship today involve meddling with trust and attention, not muzzling speech itself." (Tufekci 2018).

These concerns about the risks of open access to data should not be understood as capitulation to the current market-driven approach to scholarly publishing, nor as a call for continuation of the status quo. Instead, I hope to encourage continuation of the creative approaches to scholarship represented in this collective. I also hope the issues raised in Data Refuge will serve as a call to take greater responsibility for the systems into which scholarship flows and the structures of power and assumptions of trust (by whom, of whom) that scholarship relies on.

While plenty of participants in the Data Refuge community posited scalable technological approaches to help people trust data, none emerged that were strong enough to risk further undermining faith in science that a malicious attack might cause. Instead of focusing on technical solutions that rely on the existing systems staying roughly as they are, I would like to focus on developing networks that explore different models of trust in institutions, and that honor the values of marginalized and indigenous people. For example, in a recent paper, Stacie Williams and Jarrett Drake describe the detailed decisions they made to establish and become deserving of trust in supporting the creation of an Archive of Police Violence in Cleveland (Williams and Drake 2017). The work of Michelle Caswell and her collaborators on exploring post-custodial archives, and on engaging in radical empathy in the archives provide great models of the kind of work that I believe is necessary to establish new models of trust that might help inform new modes of sharing and relying on community information (Caswell and Cifor 2016).

Beyond seeking new ways to build trust, it has become clear that new methods are needed to help filter and contextualize publications. Our current reliance on a few for-profit companies to filter and rank what we see of the information landscape has proved to be tremendously harmful for the dissemination of facts, and has been especially dangerous to marginalized communities (Noble 2018). While the world of scholarly humanities publishing is doing somewhat better than open data or mass media, there is still a risk that without new forms of filtering and establishing quality and trustworthiness, good ideas and important scholarship will be lost in the rankings of search engines and the algorithms of social media. We need new, large scale systems to help people filter and rank the information on the open web. In our current situation, according to media theorist dana boyd, "[t]he onus is on the public to interpret what they see. To self-investigate. Since we live in a neoliberal society that prioritizes individual agency, we double down on media literacy as the 'solution' to misinformation. It's up to each of us as individuals to decide for ourselves whether or not what we're getting is true." (boyd 2018)

In closing, I'll return to the notion of Guerrilla warfare that brought this panel together. While some of our collaborators and some in the press did use the term 'Guerrilla archiving' to describe the data rescue efforts (Currie and Paris 2017), I generally did not. The work we did was indeed designed to take advantage of tactics that allow a small number of actors to resist giant state power. However,

if anything, the most direct target of these guerrilla actions in my mind was not the Trump administration. Instead, the action was designed to prompt responses by the institutions where many of us work and by communities of scholars and activists who make up these institutions. It was designed to get as many people as possible working to address the complex issues raised by the two interconnected challenges that the Data Refuge project threw into relief. The first challenge, of course, is the need for new scientific, artistic, scholarly and narrative ways of contending with the reality of global, human-made climate change. And the second challenge, as I've argued in this paper, is that our systems of establishing and signaling trustworthiness, quality, reliability and stability of information are in dire need of creative intervention as well. It is not just publishing but all of our systems for discovering, sharing, acquiring, describing and storing that scholarship that need support, maintenance, repair, and perhaps in some cases, replacement. And this work will rely on scholars, as well as expert information practitioners from a range of fields (Caswell 2016).

Closing note: The workflow established and used at Data Rescue events was designed to tackle this set of difficult issues, but needed refinement, and was retired in mid-2017. The Data Refuge project continues, led by Professor Wiggin and her colleagues and students at PPEH, who are "building a storybank to document how data lives in the world – and how it connects people, places, and non-human species." ("DataRefuge" n.d.) In addition, the set of issues raised by Data Refuge continue to inform my work and the work of many of our collaborators.

[1] At the time of this writing, we are working on un-packing and repackaging the data within Data Refuge for eventual inclusion in various Research Library Repositories.

[2] Ideally, of course, all federally produced datasets would be published in neatly packaged and more easily preservable containers, along with enough technical checks to ensure their validity (hashes, checksums, etc.) and each agency would create a periodical published inventory of datasets. But the situation we encountered with Data Refuge did not start us in anything like that situation, despite the hugely successful and important work of the employees who created and maintained data.gov. For a fuller view of this workflow, see my talk at CSVConf 2017 (Allen 2017).

References

Allen, Laurie. 2017. "Contexts and Institutions." Paper presented at *csv,conf,v3, Portland, Oregon, May 3rd 2017.* Accessed May 20, 2018. https://youtu.be/V2gwi0CRYto.

Bodó, Balázs. 2015. "Libraries in the Post - Scarcity Era." In *Copyrighting Creativity: Creative Values, Cultural Heritage Institutions and Systems of Intellectual Property*, edited by Porsdam. Routledge.

boyd, danah. 2018. "You Think You Want Media Literacy… Do You?" Data & Society: Points. March 9, 2018. https://points.datasociety.net/you-think-you-want-media-literacy-do-you-7cad6af18ec2.

Caswell, Michelle. 2016. "'The Archive' Is Not an Archives: On Acknowledging the Intellectual Contributions of Archival Studies." *Reconstruction: Studies in Contemporary Culture* 16:1 (2016) (special issue "Archives on Fire"), http://reconstruction.eserver.org/Issues/161/Caswell.shtml.

Caswell, Michelle, and Marika Cifor. 2016. "From Human Rights to Feminist Ethics: Radical Empathy in the Archives." *Archivaria* 82 (0): 23–43.

Currie, Morgan, and Britt Paris. 2017. "How the 'Guerrilla Archivists' Saved History – and Are Doing It Again under Trump." *The Conversation* (blog). February 21, 2017. https://theconversation.com/how-the-guerrilla-archivists-saved-history-and-are-doing-it-again-under-trump-72346.

"DataRefuge." n.d. PPEH Lab. Accessed May 21, 2018. http://www.ppehlab.org/datarefuge/.

"DataRescue Paths." n.d. PPEH Lab. Accessed May 20, 2018. http://www.ppehlab.org/datarefugepaths/.

"End of Term Web Archive: U.S. Government Websites." n.d. Accessed May 20, 2018. http://eotarchive.cdlib.org/.

"Environmental Data and Governance Initiative." n.d. EDGI. Accessed May 19, 2018. https://envirodatagov.org/.

Laster, Shari. 2016. "After the Election: Libraries, Librarians, and the Government - Free Government Information (FGI)." Free Government Information (FGI). November 23, 2016. https://freegovinfo.info/node/11451.

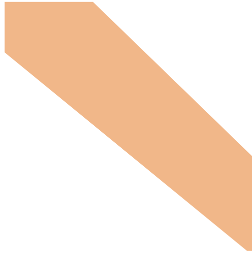
Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism.* New York: NYU Press.

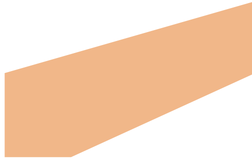
Tufekci, Zeynep. 2018. "It's the (Democracy-Poisoning) Golden Age of Free Speech." WIRED. Accessed May 20, 2018. https://www.wired.com/story/free-speech-issue-tech-turmoil-new-censorship/.

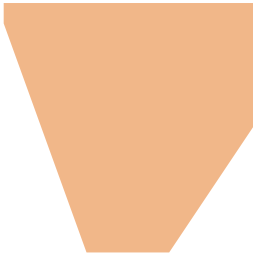
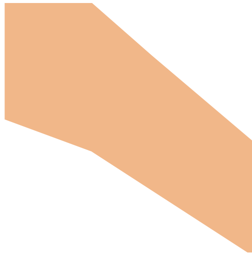
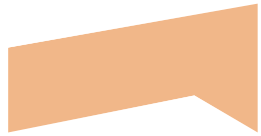
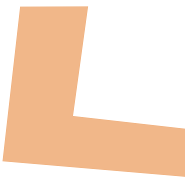
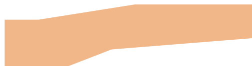
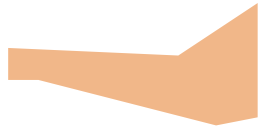
"Welcome - Data Refuge." n.d. Accessed May 20, 2018. https://www.datarefuge.org/.

Williams, Stacie M, and Jarrett Drake. 2017. "Power to the People: Documenting Police Violence in Cleveland." *Journal of Critical Library and Information Studies* 1 (2). https://doi.org/10.24242/jclis.v1i2.33.

RADICAL OPEN ACCESS
BUILDING HORIZONTAL ALLIANCES





Guerrilla
Open
Access