Final Report: Web Archive Collection

The Negative Library of Other Networks: Permacomputing & The Cozy Web

Name: Lucas Rosendall

Date: April 17th, 2026

Collection Overview

"It's a practice that encourages the maximization of hardware lifespan, minimization of energy usage, and focuses on the use of already available computational resources. It's about using computation only when it has a strengthening effect on ecosystems."

– Ville-Matias "Viznut" Heikkilä

This Negative Library of Other Networks is a collection that seeks to document the populations of the internet that have scurried away from the mainstream web's relentless pursuit of power. We can call this subject area the cozy web, where much of the content is poorly addressable, poorly searchable, and (perhaps) vulnerable to bitrot. I am interested in how, downstream of the adtech web, there have been small enclaves of communities that burrow underground, seeking out not only respite but also new protocols and foundations for the future of the internet. As most of these communities run through chat streams like Discord or Slack, the scope of this project will prioritize blogs dedicated to alternative visions for the internet. More specifically, this project will focus on cozy web or poetic web perspectives of ecology, namely permacomputing, feral web, and low-tech horizons. In a more broad sense our designated community of users are

those who are interested in a web built for futures of collective knowledge-making, for those interested in engaging with the sustainable side of the internet, and for those compelled to consider the interconnection between the world of compute and ecological understandings of the world. More narrowly this interrelated network of ideas and practices could become a critical source for historical research in a variety of disciplinary domains (some of which may still be emerging). The collection bumps up against artistic and design practices, the politics of technologies, digital methods, digital emotions, art & activism, crisis & critique, or even something like green AI and cultural production. Yet the collection isn't only for future anthropologists, social scientists, philosophers, or historians studying online movements. Permacomputing is unique in that it derives a set of design principles from whole-systems thinking. It is grounded in the physical environment. Its principles are empowered by the futurities of solar energy and electrification. Such futures involve a different relationship to consumption, or the behavioral drivers of extracting from the earth without limits. Permacomputing applies an environmental ethic to localize web ecosystems; it seeks to do so through computational restraint and solar powered servers. Currently speed and ease of retrieval, narrow modes of valuation, dominate the contemporary digital landscape. These values are constructed through environmental, cultural, and archival lenses. In this regard, the collection introduces ethical ambiguities and contradictions. The archive world is routinely stuck in asset mode, but should this collection process (and its appraisal criteria) prioritize "low-tech," or at least limited and narrowly scoped crawls

As narrow modes of valuation dominate the contemporary landscape, the 2025 edition of MediaLive digs into how value and values are constructed - environmentally, culturally, and archivally. Committed to expansive definitions of media and media arts, the exhibition champions low tech and tactile methods that resist the hyper-valuation of the latest technologies. Taken together, the works and curation make proposals for degrowth in a world seemingly fixed in asset mode.

Collection Plan

It is my intention to document the permacomputing movement much like a basketball game is documented; that is, in all its glory and misery. Yet, the collection introduces ethical ambiguities and contradictions. Currently speed and ease of retrieval, narrow modes of valuation, dominate the contemporary digital landscape. These values are constructed through environmental, cultural, and archival lenses that reify consumer behavior, rhythms, and flows. Archives too are routinely stuck in asset mode. Should this collection process (and its appraisal criteria) prioritize "low-tech," or at least limited and narrowly scoped crawls? In other words, how important is it to live out the permacomputing ethos in my methodology? It is something we will have to keep in mind and return to, especially as it concerns validity, scope, or the representativeness of the findings – the provenance, as we say in archival theory.

Overall the collection is germane to genre and domain. The culture around the so-called cozy web is driven by strands of alternative theory – anarchism, intersectional feminism, decoloniality, and degrowth. Altogether though design principles and aesthetics define the community in practice. Its computational creativity is an embrace of limits and constraints. For the most part the content in this collection remains english-heavy, but there is no telling how far a crawler will go if the crawler is not properly scoped and max depth goes unspecified. Most of

my seeds follow links to URLs that are nascent (2020 up to the present), but I imagine that the boundaries are not as fixed as I might think. All scope decisions are a reflection of the appraisal-making process (Summers, 2020) here (**Figure 1**):

Criteria	Description	Examples
Significance	Websites must be dedicated to the topics of permacomputing, low-tech sustainable computing, or the feral web. Does the site represent a key voice, position, or artifact within permacomputing discourse?	https://permacomputing.net/ https://paragraph.com/@austinwadesmith/queer-servers-and-feral-webs https://wiki.xxiivv.com/site/permacomputing.html http://viznut.fi/files/texts-en/permacomputing.html
Community Ethos	The websites should be a window into the world of sustainable computing and should therefore be written by practitioners for a wider audience; the collection is not interested in academic observation necessarily; only if there is a research community that gathers around these concepts will it be a consideration for the collection.	https://100r.co/site/permacomputing_101.html https://smallfile.ca/ https://limits.pubpub.org/

Actionable knowledge	IF the website does not mention permacomputing specifically it should provide how-tos, principles, and tools that align with low-tech infrastructure.	https://wiki.lowtechlab.org/wiki/Accueil https://solar.lowtechmagazine.com/ https://lurk.org/ https://solarprotocol.net/
Visual Culture & Content	IF the website is inspired by low tech themes and its visual aesthetic or content accords to these principles (minimalist design, sustainable, low tech computing themes).	https://128kb.eu/about/ https://pixeljoint.com/ https://smallfile.ca/
Ephemerality risk	This project is also interested in one-off essays or manifestos dedicated to writing or blogging about permacomputing. Some of these Blogs or one-off writings on websites are inactive, or were written in the early 2020s. Sites at risk of disappearing, like abandoned blogs, will also be archived as part of this collection as long as they relate to the topics at hand. How likely is the site to disappear or change substantially without notice? Small, self-hosted personal sites on non-commercial infrastructure are going to be prioritized.	http://viznut.fi/files/texts-en/permacomputing.html https://limits.pubpub.org/pub/6loh1eqi/release/1 https://permacomputingasapracticefordigitalgraphicdesign.com/ https://wiki.xxiivv.com/site/permacomputing.html https://limits.pubpub.org/pub/6loh1eqi/release/1

		https://www.sustainablewebmanifesto.com/
Permacomputing in Practice	As climate change continues to transform our environment and increasingly becomes a greater public concern, this collection would like to document the creative application of low-tech or permacomputing methodologies *in practice*.	https://www.solarpowerforartists.com/ https://meshtastic.org/ https://www.fiber-space.nl/ https://danielparnitzke.de/project/pleasureinscarcity

Figure 1: *archivalization in practice*

For this project, I used a set of tools produced by the Webrecorder project to create and make this web archive accessible. While watching the crawler run in four separate windows, I was struck by the difference between a machinic crawl or exploration of the web – browsertrix interfacing directly with the domain according to its rule set – and the more rhythmic or affective way our subjectivity impinges upon the interface. In other words, from the standpoint of our subjectivity, discovering a new link operates at the neural level. Our interaction with the computer is mediated through the screen. In this sense discovering a new link opens up a door to another world, and that door might subsequently lead to another. A crawl, on the other hand, operationalizes a different scale. In a sense it scales the domain like a mountaineer would scale a mountain, while we interface with it via screen modalities – or rather, the interface. The collection in a sense should be multi-scalar and dimensional. It should relay back to the user properties that correspond to dimensionality. Vis-à-vis the crawler – the representative seed, scope, and appraisal criteria more generally – the mountain in its dimensionality is simulated,

resurrected if you will, or brought to life in its multiscalar properties and perspectives. To crawl one seed is to scale one side (or a subset) of the mountain. The job of web archiving is to get as close to the thing as possible. It is the thing but not the thing itself, much like the bits on a TV screen reconstruct the basketball going through the hoop. Maemura and their colleagues (2018), citing Douglas (2017) refer to provence as principle (Maemura et al., 2018). This collection plan is interested in the theoretical question of what an archival aggregation represents.

Capture Settings

It's a difficult thing, capture settings. As someone who puts substantial, meticulous effort into getting things right, capture settings feel like being invited to a party you can't ever get into. Hyperlinked embeddings expand the vast oasis of other sites. A poorly configured crawl setting sends the crawler spiraling across the entire web. It's as if we are trying to tame a reactionary madman hellbent on circling the earth with Napoleon style dreams and fantasies. Crawlers are a Don Quijote style character let loose upon the narrative structure of Kafka's The Castle. Selectively culling the right or perfect materials for the collection can feel impossibly obscure – "how much material to collect in order to provide the appropriate context for archived Webpages?" Post (2017) warns us in *Building a Living, Breathing Archive.* There is also the issue of scale and the limits of technical capture, especially if you are beholden to the inadequate interface of a web archive like Browstertrix or archive-it. With that said, this collection focuses on intensity over scale – vertical craws that comprehensively capture a few sites. The frequency of the crawls in this collection are intentionally aligned with the permacomputing ethos of logistical constraint and resource limitation.

URL	Appraisal Criteria	Crawl Scope & Capture Settings	Review

Permacomputing Wiki
 	
https://permacomputing.net/
	This is a wiki page dedicated to the explanation of permacomputing and philosophy and its underlying principles. It also links heavily to other resources.	**Scope:**
		• **Crawl Scope**: Pages on Same Domain + Subdomains
		• **Crawl Start URL**: https://permacomputing.net/
		• **Max Depth In Scope**: Include Any Linked Page (“one hop out”)
		• **Exclusions**:
		  ○ - .*ikiwiki\.cgi.*
		  ○ \.*\\\?updated=\.*
		
		Crawl Limits:
		• **Max Pages**: 800
		
		Scheduling:
		• **Crawl Schedule**: 0 0 1 1,7 *
		  ○ Cron syntax for semi-annual
		
		Rationale: It captures the link context of the wiki, preserving evidence of what the community considered worth linking to.
			The tradeoff here is the crawl budget. 800 pages already seems like a suitable enough limit. But is the crawl finishing the main site before it crawls a domain a hop away?

Viznut Permacomputing Blog http://viznut.fi/texts-en/permacomputing.html	This is a blog created by permacomputing programmer and designer viznut. One of the first to discuss the concept of permacomputing online.	**Scope:** • Crawl Scope: Pages on Same Domain + Subdomains • Crawl Start URL: http://viznut.fi/texts-en/permacomputing.html • Max Depth In Scope:N/A **Crawl Limits:** • Max Pages: 800 **Scheduling:** • Crawl Schedule: @yearly	There is only one logged error where the crawl could not capture an mp3 file. There are other 404 errors but for the most part this crawl performs very well.

		Rationale: Vertical Crawls – intensive crawl that captures the domain, Interested in grabbing the entire scope of the domain so we can capture all of its history.	
Wiki - permaculture https://wiki.xxiivv.com/site/permacomputing.html		**Scope:** ● Crawl Scope: Pages on Same Domain + Subdomains ● Crawl Start URL: https://wiki.xxiivv.com/site/permacomputing.html ● Max Depth In Scope: **Crawl Limits:** ● Max Pages: 800 **Scheduling:** ● Crawl Schedule: @yearly **Rationale:**Another Wiki that I don't think is currently active. Still, it could be good to come back to it each year to see what has changed, if anything. There are many DIY programming languages that exist inside this wiki.	465 html pages captured and only 19 non-html files captured; only a few pages failed to load.
100 Rabbits - https://100r.co/site/permacomputing_101.html		**Scope:** ● Crawl Scope: Pages on Same Domain + Subdomains ● Crawl Start URL: https://100r.co/site/permacomputing_101.html ● Max Depth In Scope: **Crawl Limits:** ● Max Pages: 800 **Scheduling:** ● Crawl Schedule: @monthly **Rationale:** 100 rabbits is a research couple – one designer and one programmer – who travel the world in	This crawl runs without any failed pages.

		their small boat living out the permacomputing philosophy. We track their monthly travel website to see examples of the permacomputing ethos being implemented in practice. It really is a lifestyle. And it is not a lifestyle of individualism or isolation, but of collective action and structural reform.	
Solar Protocol https://solarprotocol.net/		**Scope:** • Crawl Scope: Pages on Same Domain + Subdomains • Crawl Start URL: https://solarprotocol.net/ • Max Depth In Scope: **Crawl Limits:** • Max Pages: 800 **Scheduling:** • Crawl Schedule: @yearly **Rationale:** Solar protocol is one of the most interesting websites I have come across. There is not any new content being developed within the domain and yet the website's aesthetic structure is constantly changing. The servers hosting the website constantly change, and how the website looks can depend on season or temperature or energy levels of the solar panels running the servers. For now we capture it for its content only; it currently is not necessary to overly crawl this website.	There are some server inconsistencies that cause a few 404 errors. Still, over 90% of the html pages are captured – we have both screenshots and the text.

Figure 2: example crawl spreadsheet

Overall there are a few sites not listed here that require more unique crawl scopes. Low-tech magazine (https://solar.lowtechmagazine.com/) would require some sort of intellectual proprietary agreement to crawl in full. In order to display all of their content in a collection, some of which is behind a pay wall, there would have to be negotiations around certain rights to present the work (This is discussed more later). Moreover, there are some crawls that are

ineffective in capturing some video and image assets, as well as the occasional hiccup with more interactive websites (https://danielparnitzke.de/project/pleasureinscarcity, https://smallfile.ca/). Some content is also not in English, but where exactly that content lies, and the languages represented, would have to be noted in the metadata. In short, crawl leakage or misrepresentation of some assets are prevalent in the collection, yet to my knowledge it is miniscule. Overall this collection is very topical in its crawl modality: all websites are adjacent to or specifically center the permacomputing ideology. Because we prioritize vertical crawls however, we miss out on cross-linked community sites, Mastodon connections, and other DIY low-tech blogs. A properly configured seed list that tries to capture the permacomputing community in all its dimensionality would maybe have to move offline, or at the very least into community discussion boards. It is out of the scope of this project, but a more fine-tuned collection set on resurrecting this small part of the web in full would have to do research outside of the crawl and capture settings of the collection.

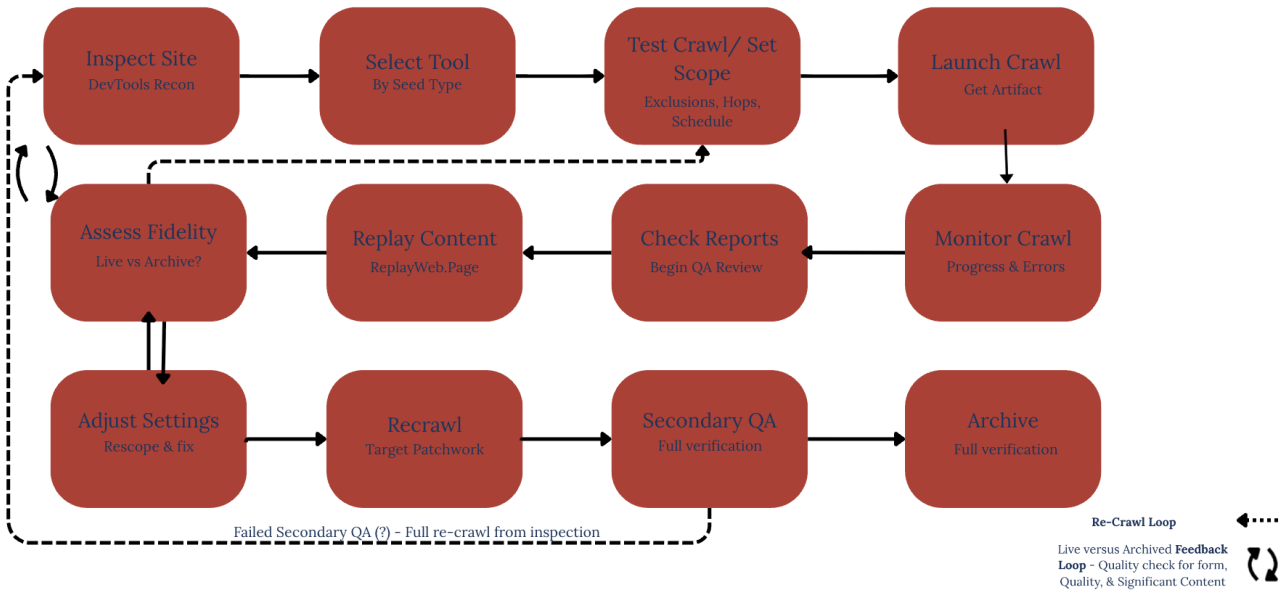


Figure 3: Capture Settings & QA workflow

Quality Assurance

As move on to Quality Assurance, **Figure 3** exhibits an iterative, generalizable workflow: preproduction (inspection → tool → test → Launch), production (the crawl), and postproduction (→ progress & errors → replay content → assess fidelity → secondary QA → Archive). In this structure each stage flows simultaneously. Importantly, a crawl is not defined by a platform like Archive-IT or BrowserTrix's Webrecorder. The archivist must select a tool by seed type, and once the artifact is obtained, assess its fidelity against the live website. This practice takes from conservatorship. If you can obtain a degree in bookbinding, it must someday be true that conservational practices meet the requirements of the web. Of course, there are questions of scale here, but with the advent of AI, which I believe will revolutionize our profession entirely, I don't see why this workflow could not be automated in the near future (and could even be done now Computer Science prescriptions I'm sure).

As for the specificities of this collection, much of the quality assurance testing was done on the https://permacomputing.net/ website, which was running one domain out into the internet. Since we can consider this website the center of our permacomputing universe, it was important to capture a scattered, horizontal, orbital understanding of what lies beyond this particular domain. What came back were links to books, github pages, journal articles, proprietary blogs like substack and medium, and discussion boards. As I inspected these sites, I moved backwards into the appraisal-making process. Were they important enough to warrant dedicated collection space and maintenance? I think it was important to register these peripheral links to see if any of them warranted their own seeds and capture settings. One of my first realizations here is that quality assurance becomes significantly less obtainable one domain out. . . and therefore links one domain out from that are just as significantly less obtainable . . . on and on . . . until your crawl is trapped in infinite recursive decay domains away from your original domain. So, I

made a conscious decision to avoid this with this collection. I think the topicality and the deep, though-provoking domain-bases made vertical crawls an easy strategy here. That said, I still encountered issues with each site. With https://permacomputing.net/ there was failed content requiring login information. In order to not waste the crawl budget and pollute the collection with low-value responses, I was able to block anything hitting ikiwiki.cgi with an exclusionary regex. With errors-first review, the check reports portion of my workflow became more concentrated on the small failures in need of further diagnosis. In Browsertrix, QA rating and screenshot comparison automated the replay workflow – I was able to sort by worst screenshot match to surface the most damaged pages in each domain.

I also ran into issues (420 errors in fact) with Solar Low-Tech Magazine. My initial thoughts were that this content was unavailable behind paywall, and that may still be the case for some content, but there were many “Page Load Failed: retry limit reached” failures. Except for about 50.5 MB captured, most of the site's content failed to load. But the broader trend of failure here showed site-wide errors, not a few broken links. I began to think about how the low-tech website is solar powered. Could it be that the crawler, hitting the domain’s pages in rapid succession, triggered page load errors? To fix this, I reduced the amount of crawls running in the browser windows to one. Then, I added page timing constraints – a generous delay between requests – so that the server would not be overwhelmed by crawler requests. Moreover, low-tech magazine website runs in Barcelona. So I set the cron schedule to hit in the afternoon UTC. I unfortunately ran out of crawling minutes because of this (Browsertrix sets limits according your subscription level 😕), but I think it was on its way to being effective, as the total size of the crawl had increased to 70 MB. Another opportunity here is to assess fidelity and change the tool. Archive-It's politeness settings, or Heritrix and wget HTTP static-based crawlers may handle this seed more gracefully.

The small file media festival website (https://smallfile.ca/) also posed its share of issues, which boil down to page not found errors. For all of the errors, the "Page not found" screen was visible for both the crawl

capture and the analysis pass, leading me to conclude that the site was already returning 404s when the crawler visited. This right here was already a tell tale sign that these broken links had nothing to do with my crawls. I am under the impression that because of the nature of the errors – paths like /programs/smallthingy/, /programs/makersforum1/, /resources/interview%20on%20carbon%20impact%20of%20streami..., and some WordPress upload paths (/wp-content/uploads/2022/07/...) – these sites have been taken down or have migrated to a new domain. If these links were incredibly necessary, the most appropriate thing to do in my eyes is to check the wayback machine. If the links exist there, then there can be proper metadata documentation dedicated to sending the user to the proper place of holdings. If the consequences were even more dire, the loss itself could be documented so that future researchers can encounter the gap with a note recording when the site was first noticed as unavailable. Otherwise this crawl is quite clean overall. The archival, full verification part of the workflow - the last step – is a very alive process in its own right, requiring digital collection maintenance and protocols for stewardship.

The Solar Protocol website (https://solarprotocol.net/) is a website that runs on a network of solar-powered servers. Some of the sunthinking subpages contain interactive art pieces that pull live data from the server network. A few pages – like live.html and sandbox.html – never seem to reach a finished loading state since they run continuous javascript animations. Our browser configurations automatically scroll to the end of pages; here, they eventually time out because the page is still executing scripts. Of course, these pages as they exist in the collection will not replay faithfully because their interactivity has not been maintained. If I wanted to improve these captures, I could increase the behavior timeout in crawl settings. In this way the crawler would wait longer before timing out. I mentioned this crawl for its interesting tensions. These interactive pages cannot be fully captured by a WARC file. The model workflow attested to above is in fact always contested by the archivists' discretion. When is fidelity worth pursuing? And what is worth it to pursue? Considering the quality of this crawl, the solar protocol artifact and its aura have already been encapsulated and made replayable for the future. If these sites were more

consequential to the collection itself, screen recording or artist collaboration would also be options. This is where archival stewardship and archivist agency walk hand in hand.

Description and Access

Field	Description & guidance	Status
Title dc:title	A Descriptive name for the collection. this can be devised by an archivist if no formal title exists.	Required
Collector dc:contributor [collector]	Who is the organization responsible for appraisal, capture, and stewardship. Distinguished from Creator element. *e.g. University of Michigan School of Information*	Required
Date dc:date	Date range of crawl activity. Use ISO 8601 (YYYY-MM-DD or YYYY/YYYY for ranges). For an ongoing collections, use open-ended range. Crawl date and content date should be distinguished. *e.g. 2025-01/ongoing*	Required
Description dc:description	A summary of the collections scope, purpose, and content. To Include: why the collection was created (overview), what types of sites are represented(plan), what is excluded(scope), and any known limitations.	Required
Subject dc:subject	Controlled vocabulary terms describing topical content in the collection.	Recommended
Genre / Form dc:format [genre]	What type of web archived content is being collected? Add secondary terms for content types present. *e.g. Wikis; Personal blogs; Community forums*	Recommended
Language dc:language	Primary language(s) of the archived content. Note any descrepencies.	Recommended
Rights dc:rights	A Statement on access conditions and re-use rights. Is online access restricted? Document if content re-use requires permission. certain rights to present their work in a public archive. Reference agreement that grants the archival institution certain rights to present their work in a public archive.	Recommended
Extent dc:format [extent]	General scope of the collection or an approximate descriptions of number of sites.	Recommended
Relation dc:relation	Links to related collections or holdings that share context with this web archive. Connect web archives to analog equivalents when deemed important.	Optional
Source of description dc:description [meta]	A Note on how the description was created: who created it, when, and from what basis; first step towards provinance transparency.	Optional

Figure 4: Data Elements for Description

The fields in the spreadsheet above (**figure 4**) make up the core data elements for our collections description. Each data element fulfills a descriptive role at multiple tiers of influence – the collection level, seed level, and document level. Because of the prospect of navigational queries, or researchers looking for a specific known site, it is important that not only the collection be described but that individual seeds and documents be described where applicable and warranted.

What is especially important is information pertaining to provenance: Why was this site selected for capture, and what is the scope of the crawl (Venlet, 2018)? In this case we have the *source of description* element brought to our attention by Dooley & Bowers (2018). This element seeks to document where descriptive information came from and to determine the context of the creation (who made the content and for what purposes?). In our examples (**Appendix A**) description, source of description, and collector all play an important role in documenting provenance level metadata. If the archive contains a transparent plan (**figure 5**), and the capture settings & QA workflow are in order, then a response to provenance capture is already in place. As part of collection maintenance, and therefore preservation and stewardship, crawl modalities and the underlying protocols become metadata. A fundamental principle of our procedures is that the provenance-level metadata is already recorded – it can exist as derivative associated metadata but it also must be embedded in these fields.

The same fields are used consistently across tiers so that metadata remains functionable across scales; moreover, consistency and standardization improves interoperability overall. However there are some existing discrepancies worth nothing that get at our metadata design choices more generally. In Appendix A, we focus our attention to the collection and seed level, but for document-level seeds the title format follows DACS conventions: *[Site Name] — [Page Title]*. Also missing from our descriptive metadata at the seed level is a creator field. Here we would follow Dooley and Bowers (2018) WAM guidelines. The collector data element may not need to be repeated at the seed-level as it already appears at the collection-level, but for safe practice institutional conditions are always important to know. At the seed level the description data element documents the sites importance to the collection. Any notable capture constraints (crawler trap exclusions, scope decisions, other limitations) will appear in the source of the

description field. This felt like one way to embed provenance metadata, but another would be to create a separate provenance-level element using Dublin Core terms. Subject terms draw from Library of Congress Subject Headings (LCSH) but this is insufficient for the web and its ever-mutating ontologies. Subject terms in this case are pulled from the pages themselves (feral computing, salvage computing, ect). As for the other fields, they document very expected additional metadata like rights, relations, language, date, and form. There is less discrepancy to be found in these fields, and for the most part they help to provide contextual access. They also improve accessibility and usability via networks of navigation and support the evaluation of archival completeness.

I would hope that Browsertrix would be able to produce some public collections gallery where the web archive can be showcased; there is always the possibility of downloading the collections WAC files and uploading them to ReplayWeb.page. Nevertheless the Browsertrix metadata schema is much thinner and is not necessarily interoperable with Dublin Core. Therefore the platform itself is not really a discovery system with vocabulary search. A finding aid becomes important in this context, especially one that is EAD encoded. Intellectual rationale for the collection, the logic of the seed list, and the crawl configuration decisions will need to be documented in narrative form. There is also the possibility of registering or bringing the collection under the control of ArchiveGrid or ArchiveSpace. But is there replay or emulation ability through those sites? One fundamental issue with the web archiving community is satisfactory access points. It is fundamentally a knowledge management problem. The collection should be able to run inside an emulator or preservation framework and should operate as a static, unchanging artifact. The question should be, what will this collection be after the internet?

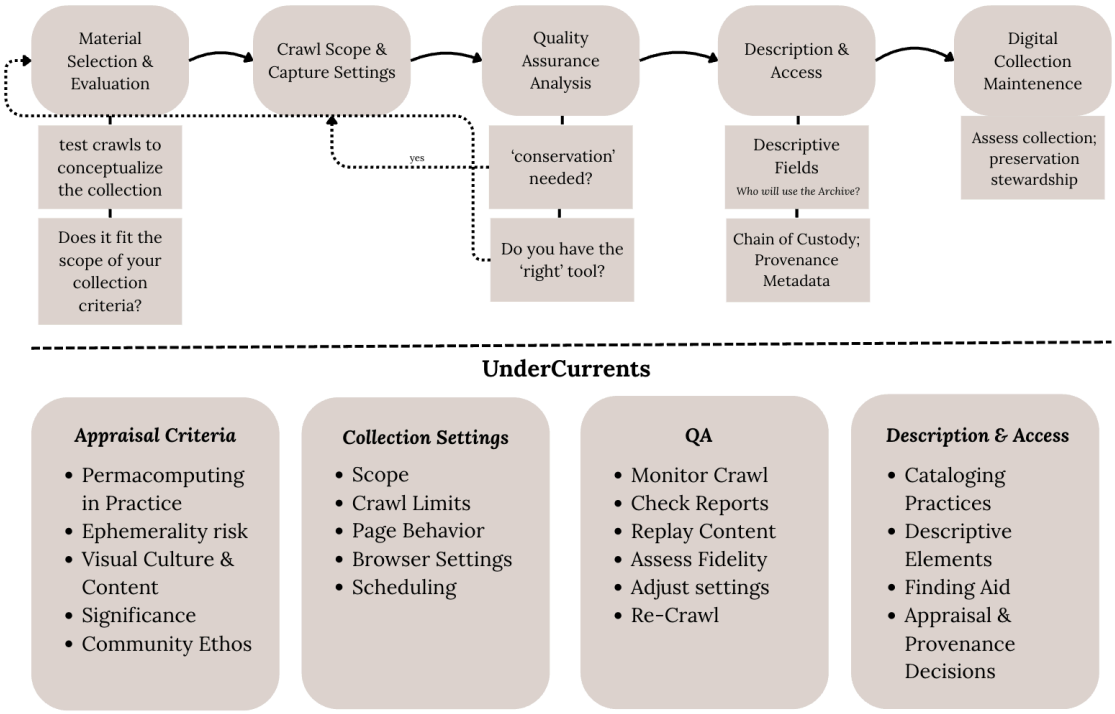


Figure 5: Web Archive LifeCycle

Intellectual Property Issues

Most of the websites would qualify as artistic works, or at least the lifestyle and the challenge to do and think otherwise is an artistic act in itself. I would think most of these sites, except those that are community forums or wikis, could be brought under some artistic agreement. This would grant the archive certain rights to present the work in a public-facing archive. Under this agreement sites like Low-Tech Magazine (https://solar.lowtechmagazine.com/), 100 Rabbits (https://100r.co/site/permacomputing_101.html), and Solar Protocol (https://solarprotocol.net/) would allow certain preservation actions to be taken without oversight or permissions. Of course, these websites would be obtained through some collection policy or formal terms of use that dictate accessibility practices and the ways in which the artifacts will be maintained with fidelity and original spirit in mind. The website(s) will only be minimally invaded and never

significantly changed unless the donor/artist comes forth with actionable requests for conservation or capture settings. Archival consent is also always an ongoing process, so works can be withdrawn or names can be removed where or when necessary. Further Intellectual property and ethical considerations would include robots.txt policies. Where it is absent or permissive, crawls are able to proceed under the Section 108 archival exception, then rights metadata notes this basis. Similarly, where a seed's robots.txt contains a disallow or a specific disallow targeting archival crawlers, that site is removed from active crawling and the disallow is documented in the Rights element of the seed's metadata record. In all cases, whether robot.txt is excluded or left in, it should be documented as provenance-level metadata. For unknowns rights holders, like Small Media Festival (https://smallfile.ca/) or the Permacomputing wiki (https://permacomputing.net/), where affirmative consent to archive can not be obtained, and a rights holder cannot be notified, then the the procedural basis for capture falls under Section 108. Should a right holder come forward asking for removal, then the collection refers to its removal policies. As always here, authority always falls to the authors and inventors of the works. Under the US constitution, exclusive rights belong to the creators. Collection policy guidelines should follow accordingly.

Appendix

Appendix A – Description of Websites for Permacomputing Collections

The Permacomputing wiki (https://permacomputing.net/), The XXIIVV wiki (https://wiki.xxiivv.com/site/permacomputing.html), and The Viznut website

(http://viznut.fi/texts-en/permacomputing.html) were all chosen for their proximity. All of them together represent the foundational seeds for the collection and the permacomputing community more generally. Each of them are similar enough but they all contain different scope parameters and providence-level metadata.

0. Collection-level Record – Permacomputing Web Archive

Title	dc:title	Permacomputing Web Archive, 2026–
Collector	dc:contributor [collector]	University of Michigan, School of Information (SI639: Web Archiving)
Date	dc:date	2026-01-01/ongoing
Description	dc:description	Permacomputing is a design practice that encourages the maximization of hardware lifespan, minimization of energy usage and focuses on the use of already available computational resources. It values maintenance and refactoring of systems to keep them efficient, instead of planned obsolescence, permacomputing practices planned longevity. It is about using computation only when it has a strengthening effect on ecosystems. This archive seeks to document the permacomputing movement and affiliated low-tech communities. The collection was created to preserve a geographically dispersed, primarily independent-web community. This collection's specific purpose is to focus on the domain of ecology, and seeks to acquire web content dedicated to permacomputing, the feral web, and low-tech domains. Seeds were selected through hand-curation to represent four overlapping strata of the movement: (1) primary philosophical hubs and foundational texts; (2) practitioner wikis and project sites; (3) community ethos; and (4) adjacent design, media arts, and low-technology organizations doing permacomputing in practice. Content types include collaboratively maintained wikis, personal essays, net

		art and engineering projects, educational resources, solar-powered websites, and community forums. Crawls conducted semi-annually (March and September) with monthly recrawls for highest-priority seeds.
Subject	dc:subject	Permacomputing; Sustainable computing; Computer science—Philosophy; frugal computing; Low-tech technology; Small web; Online communities; Internet—Social aspects; Sustainable development—Technological aspects; Net art; Design—Environmental aspects
Genre / Form	dc:format [genre]	Web archives; Community wikis; Personal websites and blogs; Essays; Net art; Project documentation websites; Educational websites; Solar-powered websites; Festival websites
Language	dc:language	en (primary); fr (Low-tech Lab Wiki); en; fr; fi (Viznut permacomputing Website)
Rights	dc:rights	No formal donor agreements in place; collection operates under Brosertrix's standard terms of service. Seeds under Creative Commons licenses noted at seed level. Access provided via Browsertrix. Content re-use subject to rights of individual site creators.
Extent	dc:format [extent]	10 seeds; approx. 8.3 GB total capture; crawl activity spanning 2026-01-01 to present
Relation	dc:relation	Related holdings: Permacomputing community forum at permacomputing.net/Community

Source of description	dc:description [meta]	Collection-level record created by El Rose, April 2026; based on capture settings and QA review workflow; crawl documentation available. Subject terms drawn from Library of Congress Subject Headings (LCSH) and Getty Art & Architecture Thesaurus (AAT). Genre/Form terms from Library of Congress Genre/Form Terms (LCGFT). Appraisal decisions available in documentation. Date formatted per ISO 8601.

1. https://permacomputing.net/

Field	DC Element	Value
Title	dc:title	Permacomputing Wiki
Collector	dc:contributor [collector]	University of Michigan School of Information, SI639
Date	dc:date	2026-04-12/ongoing
Description	dc:description	The canonical community – maintained wiki for the permacomputing movement. Defines foundational concepts, values, and principles including hardware longevity, software minimalism, and energy-aware design. Primary intellectual hub for the collection.

Subject	dc:subject	Permacomputing; Sustainable computing; Computer science—Philosophy; frugal computing; Low-tech technology; Open-source software
Genre / Form	dc:format [genre]	Community wikis
Language	dc:language	en
Rights	dc:rights	Content under open community license; Crawled in compliance with robots.txt. Archived under…
Extent	dc:format [extent]	1 crawl; 3.25 GB
Relation	dc:relation	Part of permacomputing collection
Source of description	dc:description [meta]	Metadata created by El Rose, April 2026, based on capture settings and QA review workflow; crawl documentation available. Excludes user talk pages and wiki administrative pages; ikiwiki crawler trap exclusions applied (.*ikiwiki\.cgi.*).

2. https://wiki.xxiivv.com/site/permacomputing.html

Field	DC Element	Value
Title	dc:title	XXIIVV Wiki — Permacomputing

Collector	dc:contributor [collector]	University of Michigan School of Information, SI639
Date	dc:date	2026-04-13/ongoing
Description	dc:description	Devine Lu Linvega's permacomputing page within the XXIIVV personal wiki. Linvega (Hundred Rabbits) is a practitioner of the movement; this page documents their permacomputing values and links to self-developed constrained-computing tools (Uxn, Orca).
Subject	dc:subject	Permacomputing; Uxn (virtual machine); Constraint-based design; Sustainable computing; Personal wikis
Genre / Form	dc:format [genre]	community wikis
Language	dc:language	en
Rights	dc:rights	Content © Devine Lu Linvega. Crawled in compliance with robots.txt. Archived under…
Extent	dc:format [extent]	1 crawl; 554 MB
Relation	dc:relation	Part of Permacomputing Collections Related to: https://100r.co (same creator); https://permacomputing.net

Source of description	dc:description [meta]	Metadata created by El Rose, April 2026, based on capture settings and QA review workflow; crawl documentation available; Seed scoped to /site/permacomputing.html rather than root domain to avoid capturing the full XXIIVV wiki, whose scope extends well beyond permacomputing.

3. http://viznut.fi/texts-en/permacomputing.html

Field	DC Element	Value
Title	dc:title	Permacomputing (Foundational Essay & website)
Collector	dc:contributor [collector]	University of Michigan School of Information, SI639
Date	dc:date	2026-04-13/ongoing
Description	dc:description	A foundational permacomputing essay written by Ville-Matias Heikkilä (viznut), originally published c. 2020. This text is often cited as coining the term "permacomputing." Viznut establishes its philosophical framework drawing on permaculture principles as applied to

		computing. The seed also includes examples of constrained coding & design.
Subject	dc:subject	Permacomputing; Computer science—Philosophy; Sustainable computing; Permaculture; low-tech; cozy-web
Genre / Form	dc:format [genre]	Essays; Personal blog & website
Language	dc:language	en; fi
Rights	dc:rights	Content © Ville-Matias Heikkilä. Crawled in compliance with robots.txt. Archived under . . .
Extent	dc:format [extent]	1 crawl; 175 MB
Relation	dc:relation	Foundational citation for: permacomputing.net & https://wiki.xxiivv.com/site/permacomputing.html
Source of description	dc:description [meta]	Metadata created by El Rose, April 2026, based on capture settings and QA review workflow; crawl documentation available. Seed includes English-language texts and Finnish-language texts.

References

Douglas, J. (2017). Origins and beyond: the ongoing evolution of archival ideas about provenance. In H. MacNeil & T. Eastwood (Eds.), Currents of archival thinking (2nd ed.). Santa Barbara, CA: Libraries Unlimited.

Maemura, E., Worby, N., Milligan, I., & Becker, C. (2018). If these crawls could talk: Studying and documenting web archives provenance. Journal of the Association for Information Science and Technology, 69(10), 1223–1233. https://doi.org/10.1002/asi.24048

Post, C. (2017). Building a Living, Breathing Archive: A Review of Appraisal Theories and Approaches for Web Archives. Preservation, Digital Technology & Culture, 46(2), 69–77. https://doi.org/10.1515/pdtc-2016-0031

Summers, E. (2020). Appraisal Talk in Web Archives. *Archivaria*, 70–103.

Dooley, J., & Bowers, K. (2018). *Descriptive Metadata for Web Archiving: Recommendations of the OCLC Research Library Partnership Web Archiving Metadata Working Group*. https://doi.org/10.25333/C3005C

Venlet, J., Farrell, K. S., Kim, T., O'Dell, A. J., & Dooley, J. (2018). *Descriptive Metadata for Web Archiving: Literature Review of User Needs*. https://doi.org/10.25333/C33P7Z

Wiedeman, G. (2019). *Describing Web Archives: A Computer-Assisted Approach*.