

Nathalie Fridzema* & Iris Geldermans

Excavating the Archived Web: Writing a Vernacular Web History following Media-Archaeological Principles

Abstract

This article reflects on how media-archaeological principles can inform research with web archives and the writing of cultural web history through a case study of the XS4ALL homepage collection at the National Library of the Netherlands. Created by everyday users in the 1990s and early 2000s, these homepages offer rare insight into vernacular web practices, while their retrospective, selective archiving raises fundamental questions of materiality, temporality, and source criticism. The article presents a methodological reflection organised around three media-archaeological themes: the materiality of archived websites and the techno-cultural archive; experimental web archaeology, focusing on the epistemic role of tools, software, and collaboration; and the writing of cultural web history, attentive to non-linear trajectories, everyday practices, and recurring imaginaries. It demonstrates how media archaeology provides a framework for acknowledging the partiality of web archives while affirming their necessity, and asserts that its principles are vital for the growing fields of web archive studies and digital humanities.

Keywords

Cultural Web History | Web Archives | Vernacular Web Practices | XS4ALL | Critical Archival Studies

Trefwoorden

culturele webgeschiedenis | webarchieven | alledaagse web praktijken | XS4ALL | kritische archief studies

Introduction

Media archaeology, Jussi Parikka has suggested, is a ‘travelling discipline’. Untethered to a single field or object, it provides a framework for studying complex temporalities that exceed ‘individualizing stories of heroes’. As such, ‘the media archaeological agenda includes much more than the past and the

*corresponding/first author

present – it points to archives of the future’.¹ Web archives can be seen as precisely such archives of the future. They have grown into invaluable sources, preserving early iterations of the web and enabling investigations into online cultures and media practices that continue to shape our contemporary digital reality.² In parallel, media-historical scholarship has increasingly turned toward the 1980s and 1990s, recognising these decades as pivotal for the emergence of digital technologies such as the World Wide Web.³ In this context, web archives are essential; they not only preserve traces of the past but also provide the means to re-examine the cultural trajectories of the present.

The expansion of web archive studies reflects this shift. Over the past decades, web archiving has increasingly moved from experimentation towards institutionalisation, developing more formalised processes and policies alongside forms of distributed innovation that emerge through collaboration among national archiving institutions, researchers, and other external partners.⁴ Technological and procedural advancements, including improved search capabilities and the development of metadata- and software-based infrastructures, have facilitated the growing use of large-scale and quantitative analyses of web collections.⁵ At the same time, making sense of archived websites and web data remains far from straightforward, as shown by a substantial body of scholarship that has engaged with the methodological challenges of using web archives for historical or qualitative research, addressing questions of source criticism, archival mediation, and the interpretive work required to reconstruct traces of the past meaningfully.⁶ Rather than treating archived websites as stable historical documents, this work foregrounds the instability, partiality, and infrastructural conditions of web archives and objects, reflecting on how historical narratives and practices are shaped through archival engagement.

In the Netherlands, web archiving is an established field that continues to develop in distinct ways. The Dutch state increasingly recognises digital heritage as significant, and regulatory frameworks aimed at stimulating sustainable archiving practices are currently being developed.⁷ At the same time, scholarship on Dutch web history is expanding alongside the development of methodological approaches for working with Dutch web collections.⁸ More specifically, a web archaeological orientation is rooted in the Dutch context, most visibly in the 2019 special issue of *TMG Journal for Media History* on web archaeology and in subsequent studies, including Robert Jansma’s software-archaeological exploration of *De Digitale Stad* and Kees Teszelszky’s pioneering work on the selection and harvesting of Dutch web incunables at the National Library of the Netherlands (Koninklijke Bibliotheek, or KB).⁹ Building on this work and operating within the current web archive

landscape, the present article contributes an updated methodological investigation centred on a singular vernacular web collection, foregrounding the research process itself as an analytical site.

This article is based on research done with the XS4ALL homepage collection, stored at the KB.¹⁰ Founded in 1993, XS4ALL (read as ‘Access for All’) was among the first public internet service providers (ISPs) in the Netherlands.¹¹ True to its name, the organisation sought to make the internet accessible beyond the confines of academia and professional circles, becoming a key player in shaping the Dutch web at a moment when its cultural and social implications were only beginning to take shape.¹² Every XS4ALL subscriber was given space to create a personal homepage, without standardised templates, established rules, or fixed expectations. As a result, the collection captures the practices of a first generation of everyday web users – those who came after the pioneers and computer hobbyists – as they experimented with how to represent themselves and what to do online. It offers rare insight into vernacular creativity and expression on the web, making the homepage collection an especially rich source for digital cultural history. Its significance is further underscored by UNESCO’s recognition of it as digital heritage.¹³

Media archaeology offers a productive framework to explore this kind of collection. Theoretically, its attention to the entanglement of cultural practices with technological infrastructures, its sensitivity to non-linear temporalities, and its focus on everyday uses of media make it particularly suited for early web archives such as XS4ALL. Specifically, more recent media-archaeological scholarship offers practical tools for critically engaging with web archives by drawing attention to the epistemological layers at play when interacting with this type of archive and object, thereby shaping what we can know. Consequently, this article has a primarily methodological objective, reflecting on the research process through the lens of media-archaeological principles and studying how they can inform the writing of a cultural history of the web using early web collections. In doing so, the article builds on other reflective accounts from the field, such as those of Megan Sapnar Ankerson or Sophie Gebeil and Jérôme Thièvre, whose emphasis did not lie on presenting a standardised methodological procedure, but on tracing how historical knowledge is shaped through archival engagement, technical experimentation, and interpretive decision-making.¹⁴

Reflecting on our work with the XS4ALL homepage collection, this article positions media archaeology as an essential framework for working with archived vernacular web pages, whose form, temporality, and availability are shaped by archival and technical infrastructures. The following

sections develop this argument through three media-archaeological themes: (1) the materiality of web archives, (2) experimental web archaeology, and (3) writing a cultural web history.

The Materiality of Web Archives

Media archaeology's focus on materiality is not straightforward, as scholars have studied various materialities. According to Parikka, it is possible to distinguish among 1) materialities of cultural practice, 2) materialities of materials, and 3) materialities of technologies.¹⁵ For this first section, the final framework proved most useful, being grounded in Kittler's push for media-specificity. 'This has meant digging into how technologies work, and finding structures of power through a technological analysis. Media archaeology is, for theorists such as Ernst, reverse-engineering'.¹⁶ Following this orientation towards websites expands the analytical lens to study more than what appears on screen. It includes the website's modalities (text, source code, MIME types, aesthetics) and its operational logic, particular to the technology of the World Wide Web, grounded in a specific historical period. As Parikka argues: '[m]edia archaeology *has* to be media specific'.¹⁷ This statement resonates with Niels Brügger's discussions of a web-minded approach as a distinct operational lens that emphasises attentiveness to the processes of web capture, the materiality of archived pages, and the infrastructural contingencies that shape their preservation.¹⁸ Here, he makes an important distinction: a website and an archived website are not the same, nor is one simply a copy of the other. Rather, the act of archiving produces a unique version, what Brügger calls 'a document of the web', made from raw materials but transformed into a new source.¹⁹ The website is not taken out of circulation intact but is *reconstructed*. The archival system and its processes, therefore, shape what the researcher encounters. This observation is not new, as Brügger's call for a new 'website philology' dates back to 2008.²⁰ Circling back to media archaeology's concern with materiality and medium-specificity, we must recognise that *archived* websites are our object of study.

Parikka has already positioned the *archive* as a crucial 'site' of memory. He writes that we should bring 'media archaeology into proximity with the archive as a key site of digital software culture, as well as (...) bringing media studies ideas into proximity with key non-academic institutions involved in cultural heritage in the digital age'.²¹ Drawing on Foucault, Parikka shows that the archive is not solely a repository of sources but also a site of knowledge production in itself, hidden in the machine: 'the software, the hardware, the protocols, and platforms which form the visibility, the audibility, the

statements of what is'.²² Extending this to archived websites means recognising that the archive (the archival infrastructure) becomes a place from which knowledge can be derived. Historical studies of the web must therefore acknowledge three epistemological materialities: the past web, the archived constructions, and the archival environment that shapes what survives, where it resides, and how it can be studied.

Here, Wolfgang Ernst's notion of the techno-cultural archive becomes helpful. It frames the archive as simultaneously a cultural product and a technological system. As he puts it, '[m]edia archaeology is primarily interested in the nondiscursive infrastructure and (hidden) programs of media. Thus, it turns from the historiographical to the techno-archival (archaeographical) mode, describing the nondiscursive practices specific to the element of the techno-cultural archive'.²³ The interplay between human curatorial choices and machine processes thus defines the techno-cultural archive.

The XS4ALL Homepage Collection as Techno-Cultural Archive

Approaching a web archive through the techno-archival mode means examining how it stores, orders, and makes sources retrievable. The XS4ALL homepage collection, housed within the KB's web collection, serves as a particularly revealing case.²⁴ Compared with the Internet Archive or other web collections at national libraries, the XS4ALL collection is distinctive for two reasons. First, the homepages were reconstructed from a historic yet still-active domain. Many of these sites, created in the 1990s, had remained online – largely forgotten and untouched – until the provider KPN, which had acquired XS4ALL in 1998, announced the company's gradual shutdown.²⁵ As a result, many homepages linger in a state of abandonment, more like online fossils than living sites, and are now steadily disappearing from the internet (see figure 1). Recognised as part of Dutch digital heritage, the KB intervened in 2019 to preserve them.²⁶ Unlike most contemporary web archives, which capture the web in real time and date materials precisely to the moment of harvest, the XS4ALL collection was thus assembled retrospectively, complicating questions of authenticity and carrying important implications for source criticism.

Secondly, the XS4ALL collection is both selective and manually curated. This selectivity stems from copyright restrictions and, in the Dutch context, the absence of a legal deposit law.²⁷ As a result, the KB could not harvest the entire XS4ALL domain but instead relied on a manually compiled list of URLs. Inevitably, multiple layers of subjectivity are inscribed in the collection, from the recognition of

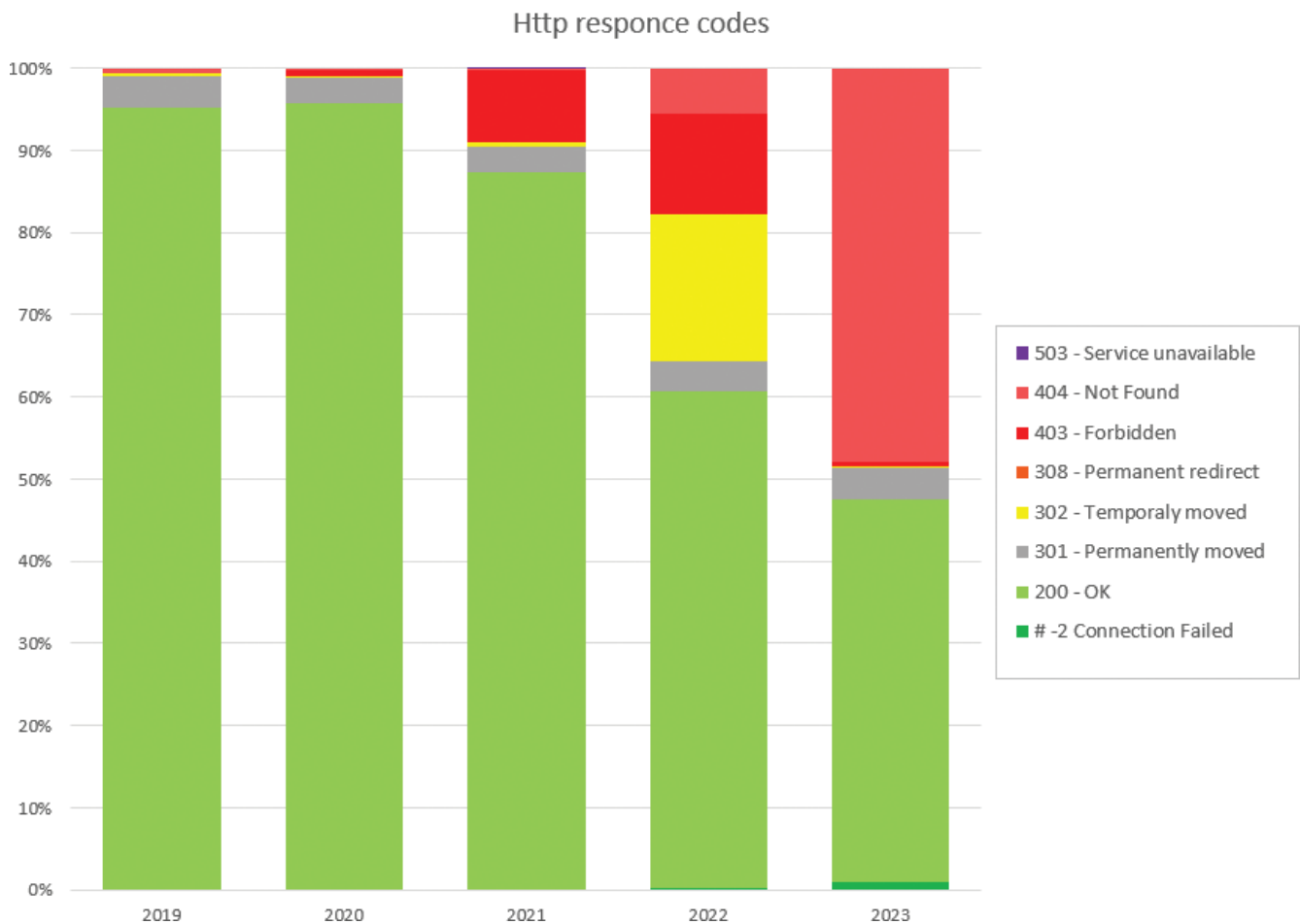


Figure 1. Overview of percentages of HTTP response codes found in approx. 19,000 XS4ALL homepages on the live web. Green indicates that the website still exists, indicating a steady decline over time²⁸

XS4ALL as a site of heritage worth preserving, to the curators' decisions about which pages held historical value, and finally to the ways these choices shaped how the homepages can now be interpreted.

We can learn more about the curatorial process behind the XS4ALL homepages in its collection description, which is openly accessible.²⁹ Here, we see that the national library's selection policy for digital collections emphasises representativeness, focusing primarily on Dutch language, culture, politics, history, and society. The curators explicitly describe their cultural-historical valuation as a bridge between societal and scholarly value, grounding their choices in both heritage preservation and academic relevance.³⁰ Furthermore, the collection description stresses that the selection process is always partial; due to opt-out procedures and manual work, not every homepage can be traced or

archived, rendering the collection inherently non-representative and making careful valuation essential.³¹

The selection process followed five main criteria: historical value (with age as the most decisive factor), aesthetic and memorial qualities, physical integrity, content value (rarity, informativeness, representativeness), and archiveability.³² These criteria resulted in striking inclusions and exclusions. Some sites were selected as cultural masterpieces, while others were excluded for being too fragmentary or generic. Yet aesthetic judgement also played a role; seemingly ‘ugly’ pages, filled with GIFs or Comic Sans, were retained for their cultural or memorial significance. Virtual art and music sites, web design experiments, and link lists were almost always preserved as they captured digital cultural practices at risk of vanishing for good. Notably, however, websites with a lot of pictures were mostly not preserved due to privacy concerns, including homepages about holidays, house renovations, or personal diaries. Websites with advertisements or about genealogy are also limited, the latter specifically because these sites are often ‘too big’ and ‘only matter to the family’, according to the collection description.³³ As such, descriptions of the collection’s creation not only reveal the content and its reasoning but also the gaps and silences to be taken into account when conducting research with this web archive.

From a technical perspective, the XS4ALL homepage collection reflects the KB’s harvesting practice of capturing one website at a time, once a year, through the Web Curator Tool and the Hetrix crawler.³⁴ This method has the advantage of producing clearly delineated harvests, each assigned its own identifier and WARC file, which makes them relatively straightforward to separate and select for research. At the same time, the single-site focus risks losing broader context. In the case of XS4ALL, for instance, the domain structure was changed in 2011, when addresses in the form of *xs4all.nl/~website* migrated to *website.home.xs4all.nl*. Traces of this shift remain in the pages themselves, yet the KB’s strictly defined harvest settings exclude such ‘old’ links when capturing the *home.xs4all.nl* domain (see figures 2 and 3). Unless curators deliberately account for this during selection or quality assurance, these fragments disappear from the archived collection, highlighting how technical parameters directly shape what survives of the cultural record.

The XS4ALL collection, understood as a techno-cultural archive, is shaped not only by technological contingencies but also by processes of cultural valuation, curatorial choice, and legal frameworks. The following section considers what this means for the historian’s preparatory work,

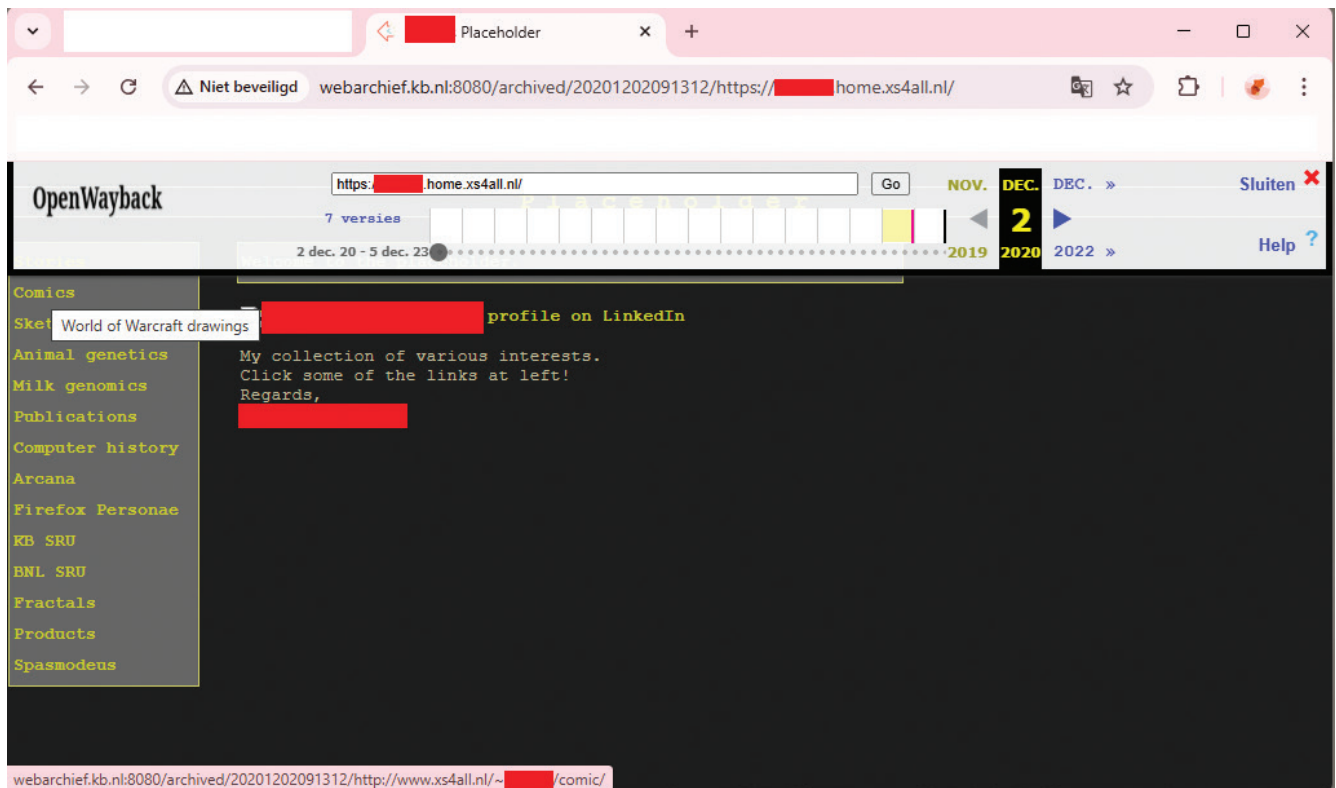


Figure 2. Example of a 'broken' website: *home.xs4all.nl*. All hyperlinks in the menu on the right are structured following the format *xs4all.nl/~website*

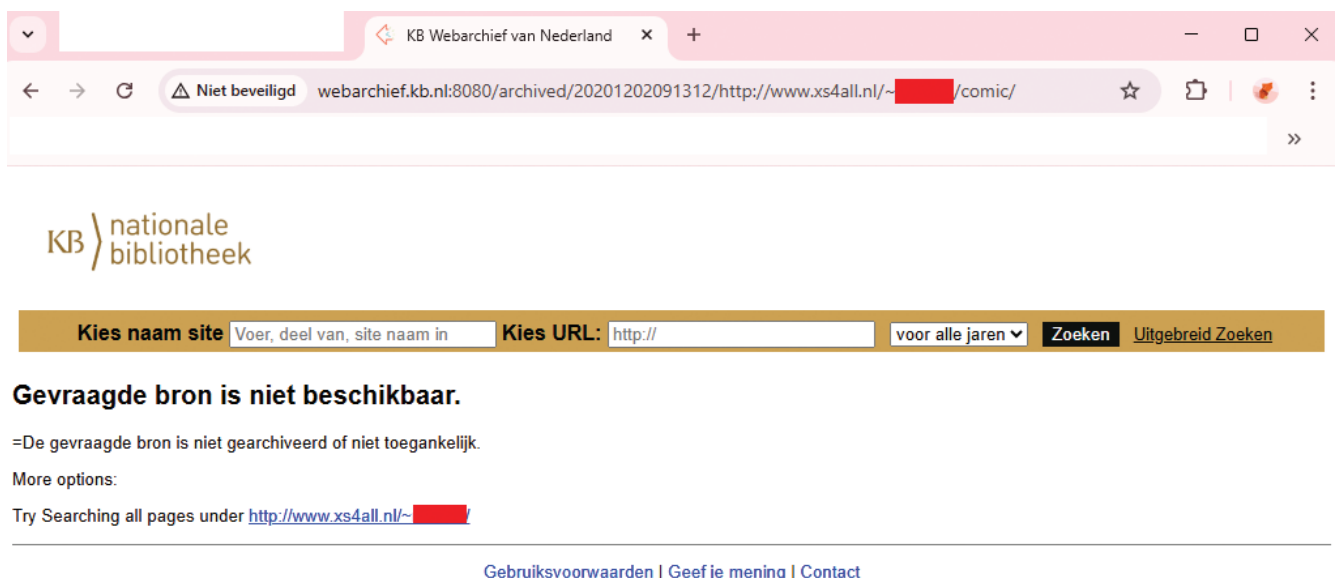


Figure 3. XS4ALL homepages with an 'old' URL are not being recognised by the crawler, resulting in these parts missing from the archived website

including how the collection's subsets can be transformed into researchable data samples, and how issues of chronology and source criticism must be addressed.

Preliminary Selection and Preparing Samples

Starting web archive research can feel overwhelming. In a physical archive, a researcher can walk into the reading room, see the shelves and boxes, and quickly form a sense of scale, order, and content. Skimming materials helps to build familiarity and intuition. By contrast, most web archives do not offer such an immediate, visual overview (yet). What confronts the researcher is often no more than a long list of URLs with metadata. In contrast, digitised archives of newspapers, books, or television programmes often provide powerful search interfaces with keyword queries and rich metadata fields.³⁵

Web archives are more differentiated: the Internet Archive's Wayback Machine does not have a full-text search engine and works best starting with a specific URL, while other tools, such as SolrWayback, allow keyword searches across various metadata fields.³⁶ Yet even the latter poses challenges when used as a starting point, due to the multimodal nature of web data and the software's ranking logic – points we return to later. Thus, where analogue or digitised collections allow orientation through *browsing* more easily, web archives, ironically, demand deliberate preparatory work. The historian must first identify which URLs are relevant for their research objective and how to navigate the archive's affordances.

In the case of the XS4ALL homepages, preparatory work was essential. Dutch legal restrictions mean that the collection cannot be accessed remotely, so researchers must make the most of the limited time they are granted in the KB reading room. Yet even if this were not the case, preliminary selection proved crucial to making the material manageable in the SolrWayback environment and, later, for close reading. SolrWayback, the KB's software of choice for research purposes with their web collections, is an open-source exploration tool developed by the Royal Danish Library that allows researchers to search and analyse web archives by indexing content extracted from WARC files, including HTML text and embedded media, facilitating qualitative and exploratory research.³⁷

Our starting point was a list of 3,276 URLs with several metadata fields, the most important being the curator's notes and the title of each page (as defined by the homepage builder in the source code). Drawing on prior research and the relevant literature, the collection description,³⁸ and by glancing at websites still online or available via the Internet Archive, we grouped the collection into

two thematic subsets: professional websites and homepages related to everyday cultures. The private/public, or domestic/professional juxtaposition, has been proposed by Ankerson, who also foregrounds the fluidity of these categories. Often, professional websites included personal elements, and vice versa, reflecting the open attitude of the early vernacular web.³⁹ This means, for our subsets, that some homepages are part of both, and that the ‘professional’ subset does not solely include company websites based on more generic, contemporary web design practices. The two subsets were not predetermined but emerged through an abductive exploration; this iterative approach allows researchers to apply existing theories while remaining open to revision in light of new findings.⁴⁰ The collection description was especially helpful at first, as it already identified common types of homepages or elements present in the collection (genealogy sites, link lists, CVs, or picture albums, for example). Especially for a collection that had not yet been studied, this approach was invaluable. It provided an initial orientation to early web practices – a kind of datafied skeleton that could then be fleshed out through deeper exploration of the collection’s specificities, nuances, and outliers at a later stage.

Once the two subsets had been defined, the next step was to prepare them for exploration in SolrWayback. This process involved deduplication, since each URL had been harvested annually across several years. To search the entire collections and subsets in SolrWayback, we selected the versions with the largest file sizes, assuming these captures contained the most data and were therefore least fragmented. For qualitative close reading, however, we drew on the earliest available harvests, reasoning that they were closest to the sites’ original form. Using different harvests for different purposes underscores how web data must be handled in relation to a study’s specific objectives, and how the act of ‘building’ the archive itself is a techno-cultural intervention.

For historical studies, determining dates and authenticity through source criticism is a crucial step. The XS4ALL collection was archived retrospectively, meaning the moment of capture does not necessarily reflect when a site was created or last updated, which complicates this step. The KB formally dates the collection to 1993–2001,⁴¹ but this timeframe relies on a patchwork of evidence: visible time stamps, source-code traces, hidden metadata, file sizes, and period-specific software or terminology.⁴² Even with such curatorial tactics, chronological certainty remains ambiguous. Also, it remains a manual and intensive labour process, resulting in some homepages that were clearly edited after 2001 still appearing in the collection. To refine our dataset, we experimented with link analysis of outgoing links as an additional digital forensics technique.⁴³ As some websites came into existence

after our period of interest – such as social media platforms like YouTube, which dates back to 2005 – domains that link to these platforms can be classified as more recent and removed from this study’s dataset altogether. In figure 4, one can see a visualisation of outgoing links revealing a cluster linking to *youtube.com*.

This difficulty in assigning stable dates illustrates why a media-archaeological perspective on materiality and the archive is indispensable. Retrospective web archives are layered and partial, and they cannot be reduced to a simple sequence of snapshots or to quantitative measures of ‘completeness’. Working with XS4ALL as a techno-cultural archive requires moving back and forth among raw data, metadata, software affordances, and close reading. Such preparatory work is foundational; it determines how the collection can be known and made meaningful for historical analysis.

Experimental Web Archaeology

After the preparatory work and the construction of subsets, the researcher can move on to studying the websites. Here, the article’s second relevant media-archaeological theme, experimental web archaeology, becomes instrumental. In *Doing Experimental Media Archaeology: Theory*, Fickers and Van

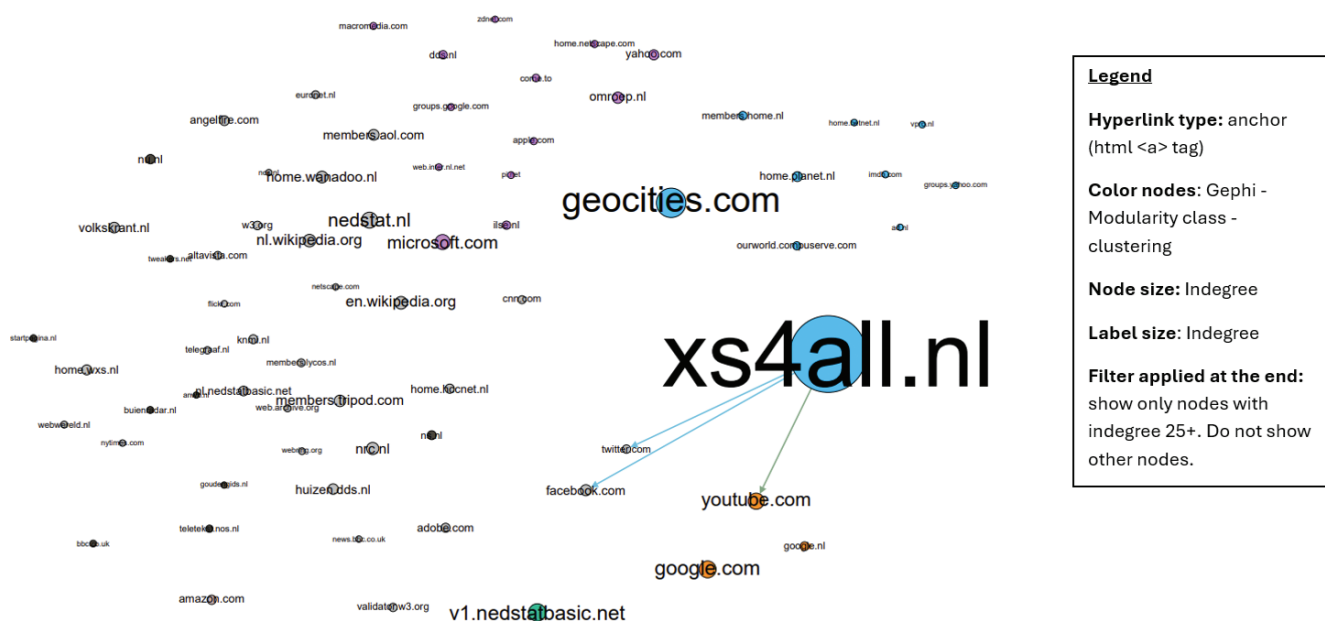


Figure 4. Network visualisation of outgoing links (everyday subset)

den Oever argue that actively working with old media technologies leads to a more adequate understanding of past media.⁴⁴ Through what they call historical re-enactment, ‘scholars experience instead of intellectually appropriate media practices’, producing tacit knowledge as well as ‘new insights in the sense of time and temporality inscribed in the materiality of media technologies’.⁴⁵ For histories that rely on vernacular sources, such as the XS4ALL homepage collection, this approach is particularly relevant, as it enables one to study appropriation strategies and shifts in the actual use of media technologies.⁴⁶ But what exactly does this approach mean?

Experimental media archaeology foregrounds the lab environment, for example, the Media Archaeology Lab.⁴⁷ In this space, media archaeology becomes a process of trial and error, creative thinking, and theory-making, where experimentation is carried out with machines. In their approach, which they term experimental media archaeology, Fickers and Van den Oever emphasise the exploratory and heuristic nature of hands-on experimentation, which requires careful self-documentation and reflexivity. As they stress:

Documenting the process of experimentation is therefore of utmost importance for the historian, as this documentation is not only a constitutive element in the construction of the ‘epistemic thing’, but also crucial ‘data’ for the later analysis and interpretation of the ‘traces’ of experimental tinkering. In fact, only the protocols turn the ephemeral traces into durable data.⁴⁸

The historian thus becomes the experimenter; rather than simply analysing past media technologies from a distance, they co-produce knowledge in interaction with the historical objects under study. As a methodological approach, Fickers and Van den Oever describe experimental media archaeology as driven by a desire to produce ‘new historical, ethnographic, and empirical knowledge about past user practices and media experiences’.⁴⁹ Media-historical sources can therefore be understood as epistemic objects; they are not fixed, but shift in meaning within the knowledge-production process. The various interpretations evolve depending on the tools, questions, or ethnographic frameworks applied. This notion is especially useful when reflecting on how objects of inquiry transform over time in response to research practices.

Several of these principles proved useful when working with the XS4ALL collection, beginning with the usage of an actual laboratory environment. Due to copyright restrictions and privacy rules,

the homepages are not accessible outside the library; researchers must instead work in the KB_Datalab, a secured space that provides both access to the collection and a range of computational tools.⁵⁰ In practice, however, the vision of the ‘historian-as-experimenter’ cannot operate in isolation, as fruitful experimentation requires collaboration among researchers, web archivists, and/or data specialists. Within this collective setting, we were able to tinker and play with the collection as the next phase of the research project.

Experimenting with XS4ALL

While experimental media archaeology often emphasises tinkering with the artefact itself, working with archived web material requires an additional layer of experimentation, engaging with the tools and software infrastructures that mediate access to the collection. As mentioned earlier, the primary tool for navigating archived web collections at the KB is SolrWayback. Its advantage lies in enabling keyword searches across the collection and, following our preparatory work, within the two thematic subsets. For a cultural-historical analysis, our ultimate aim was a close reading of individual homepages. Yet SolrWayback offered a crucial intermediary step: a way to gain a macro-level perspective on the collection, identify patterns, and detect anomalies beyond the predefined themes. In this sense, experimentation largely consisted of ‘playing’ with the tool, learning its affordances and limitations in the process.

While SolrWayback provides several built-in computational functions, these proved of limited use for a collection like XS4ALL. Date-based visualisations are based on harvest dates, which, in a retrospectively archived collection, reveal little beyond when sites were captured. Similarly, functions such as word clouds or link analyses operate at the level of the entire domain (*xs4all.nl*), rather than individual sites or curated subsets, thereby obscuring the granularity we require. As link analysis is an integral part of web research, the following sections will elaborate more on its exploratory function and the interplay between the XS4ALL collection and the SolrWayback functionality.

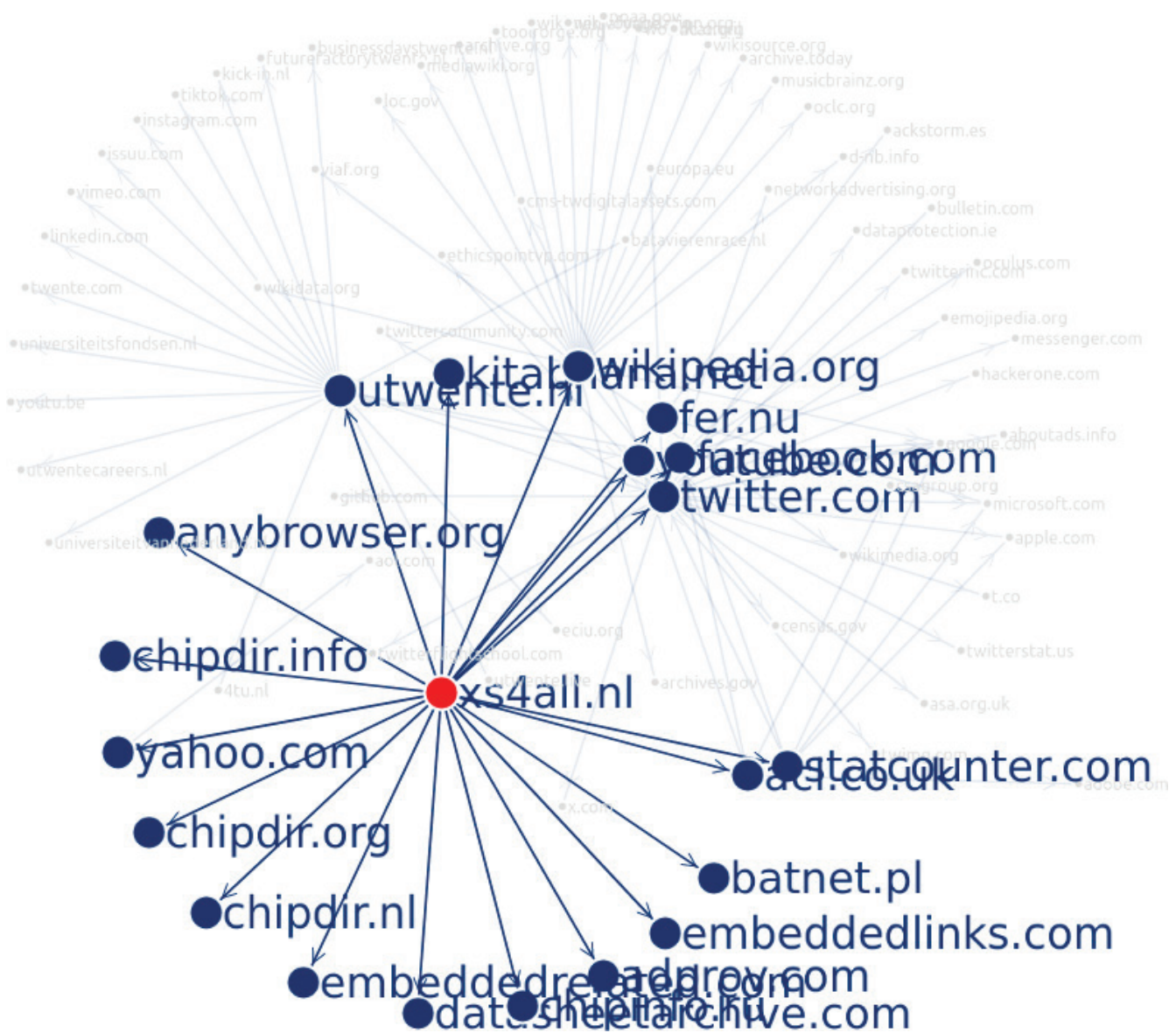
Network-based explorations and visualisations are a familiar entry point in web research, as linking is constitutive of the web’s structure and navigation. At the same time, such analyses require historical sensitivity to the social and technical specificity of linking practices, since the hyperlink is not a fixed or timeless feature. As Anne Helmond has shown, the function and meaning of hyperlinks shift across the web’s history, moving from navigation and reference to reputation and valuation.⁵¹ Situating the XS4ALL homepage collection within this framework places it firmly in a pre-platform

linking culture, where hyperlinks primarily operated as relational and navigational devices rather than as indicators of visibility or value. The linking structures visible in the collection thus reflect a different logic of connection from that of later web environments shaped by algorithmic ranking and platform architectures. Approached in this way, link analysis and visualisation can serve as an experimental starting point – a means of rendering an otherwise abstract network of homepages visible, warranting closer historical reading.

In theory, hyperlink graphs offer a powerful way to visualise such structures; they can highlight clusters, central nodes, or peripheral sites, and suggest potential points of entry for more detailed analysis. In practice, however, their value depends heavily on how links are processed and displayed through particular software. With XS4ALL, the SolrWayback implementation was of little help. Because of the software's preprocessing, all homepages were collapsed into a single *xs4all.nl* node (see figure 5), erasing the distinctions of thousands of individual sites. While documentation for the indexing process exists, it is scattered across technical GitHub pages and remains opaque to non-specialist users.⁵² This is not unique to SolrWayback. Across the board, link graphs tend to conceal more than they reveal, rarely specifying which types of hyperlinks are included or how preprocessing shapes the outcome.

For historical analysis, this opacity matters. A researcher wants to know not only which sites are linked together but also which remain unconnected. In the professional subset, for instance, more than a quarter of sites contained no external links at all – a striking finding that the SolrWayback visualisation rendered invisible. Equally important is the question of *what kinds* of links are represented. In the XS4ALL collection, most were simple anchor links embedded in page text.⁵³ Still, 10% were counters, scripts, or image links, which are technically hyperlinks but serve very different purposes (see figure 6). Because they often pointed to the same external services, such as popular web counters, they could distort the graph by inflating certain nodes, suggesting centrality where none existed in practice.

Our experiments with KB's own link-extraction scripts confirmed these distortions.⁵⁴ When applied to the professional subset, the processed results suggested that *xs4all.nl* was the overwhelmingly dominant reference point (see figure 7). At first glance, this might imply that homepage creators regularly link back to their provider. However, returning to the raw data told a different story: only nine sites were actually linked to the *xs4all.nl* domain. The majority of these supposed self-references were in fact artefacts of the 2011 domain migration, when old tilde



addresses (*xs4all.nl/~username*) were retrospectively collapsed into the new *home.xs4all.nl* structure. What appeared to be a massive node was thus the result of preprocessing, not user practice.

Another affordance of SolrWayback is its search functionality across metadata fields; queries return not just websites but distinct data objects, such as textual fragments, code snippets, or image files. A search for *under construction*, for instance, yielded not only written text but also HTML elements and graphics with that filename. Playing around with this function revealed that the tool itself mediates the research process, creating a feedback loop when used for historical inquiries.

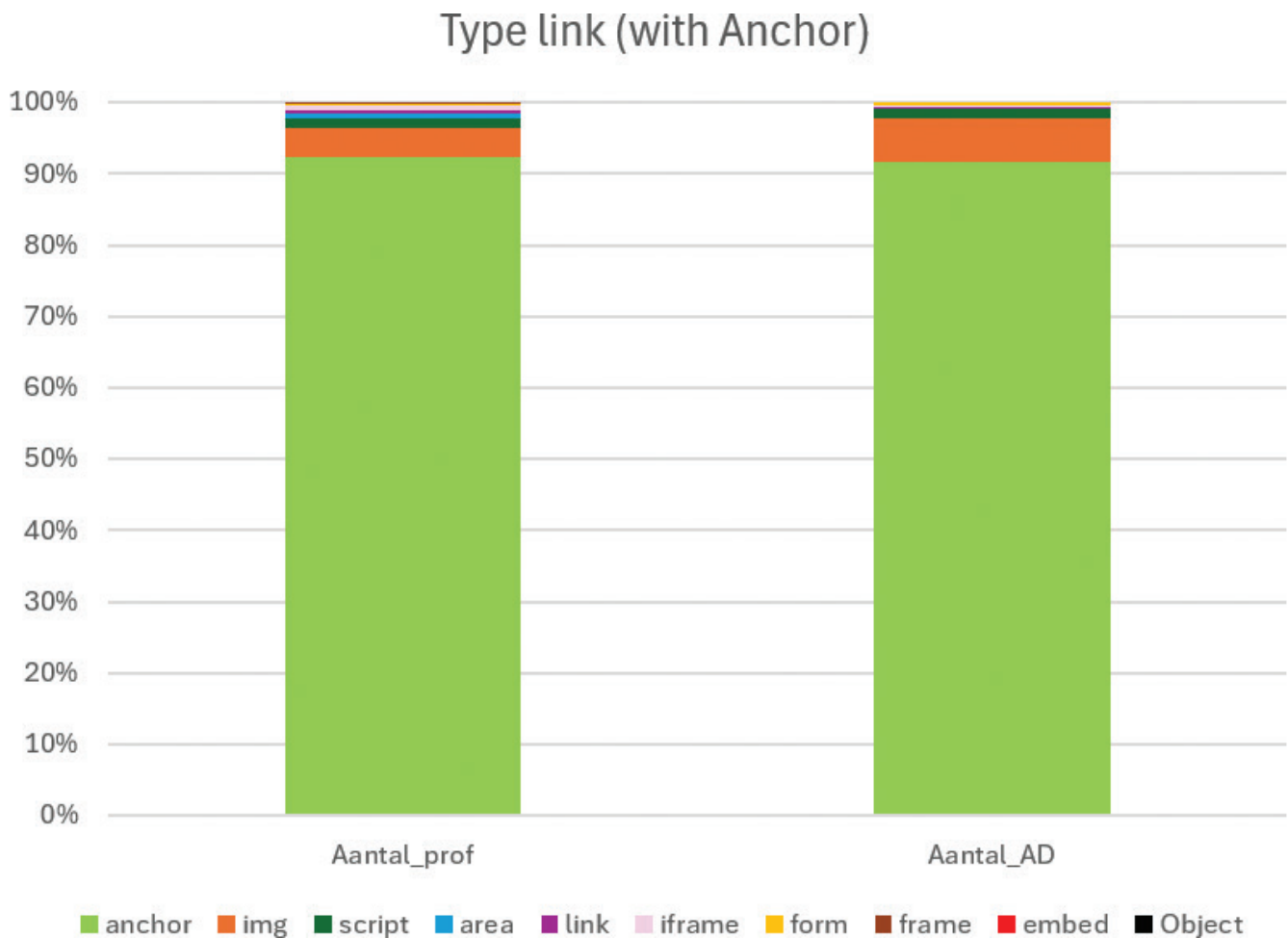


Figure 6. Types of hyperlinks in professional and everyday subsets

While searching the collection, one simultaneously learns how the tool interprets and ranks data, and how this shapes what is retrievable in the first place.

A clear example is SolrWayback's relevance ranking, or how search results are ordered. Through iterative testing, it became apparent that when a query matched terms placed behind the HTML tag *keywords* in a site's source code, that page would be ranked higher. In the context of the XS4ALL homepages, however, the use of this HTML field was likely not standardised, and amateur web makers often interpreted it differently. A search for *Groningen*, for example, might yield a personal homepage where the webmaster had listed *Groningen* under *keywords* as a personal characteristic, even though the site itself had nothing to do with the city. Lower in the results, by contrast, appeared a commercial website of a company actually based in Groningen. The outcome is not wrong, but it underscores how search results reflect both the idiosyncratic practices of amateur users and the



Figure 7. Graph visualisation of outgoing links (professional subset) using Gephi

ranking logic of the retrieval tool. For this reason, a web-minded approach is essential; SolrWayback does not provide neutral access to content but interacts with the historical material in ways shaped by past conventions of web design.

Eventually, these insights from experimenting with XS4ALL and SolrWayback helped us gain a better grasp of the collection and the tool, enabling us to extract specific URLs through queries that extended beyond the two initial subsets (*sitemap*, *guestbook*, *under construction*) for use at the close-reading stage. The link analysis experiments revealed the popularity of the NedStat visitor counter and that the homepages have a significant amount of outgoing links but minimal incoming links (from other archived homepages).⁵⁵ Specifically, the everyday subset had a vast amount of outgoing links, confirming and visualising the practice of building link lists.

Still, the epistemic implications of working with such tools become clearer when comparing XS4ALL with contemporary web archives. Unlike today's institutional or corporate sites, which are built according to broadly standardised industry formats, amateur homepages in the 1990s often

reflected individual interpretations of HTML and metadata. SolrWayback's relevance ranking thus varies by collection. With XS4ALL, the variability of amateur coding practices comes to the fore, making the tool not only a means of retrieval but also an indirect window into historical user practices. Our experimental approach to SolrWayback's search function and link analysis demonstrates the current software's bias towards contemporary web collections. Regarding the latter, built-in visualisation algorithms (also in Gephi) are better equipped to show modern web networks, which tend to be more vertically interlinked than the early web's rather horizontal, or flat, linking patterns. In this sense, the XS4ALL collection itself becomes an epistemic object. It is not a stable source waiting to be discovered, but rather has contours that shift in response to tools, queries, and interpretive choices.

The very act of experimenting with SolrWayback co-constructs what the collection is and how it can be known. Crucially, this process also highlighted what Wanda Strauven has described as the laboratory in its discursive sense; research unfolded not only through technical experimentation but also through dialogue between the historian and web archivist, translating historical questions into data operations and vice versa.⁵⁶ This collaborative negotiation resonates with the broader field of digital humanities (DH), where such translational labour between disciplinary languages and methods is a defining practice. In this way, our work with SolrWayback illustrates how experimental media archaeology, as a principle, can serve DH by producing knowledge through the interplay of archives, tools, and collaborative expertise. This approach is particularly relevant in DH fields that are relatively young, where data practices are not yet standardised, whether for legal or technical reasons, and collections often remain fragmented. For historical studies in particular, such dialogues among the archive, the tool, the researcher, and past user practices demonstrate how an experimental media-archaeological approach can benefit from the thorough application of source criticism.

Writing a Cultural Web History

The previous sections showed how media-archaeological principles helped us prepare the XS4ALL homepage collection and experiment with the SolrWayback, giving us a better sense of how to interact with the archive. These steps provide a fruitful foundation for exploring the homepages more qualitatively and making sense of the everyday practices of a first generation of users. Media

archaeology offers several perspectives that align with this objective by shifting attention away from invention or technical sophistication and towards ordinary use – what people actually did with the web. As Fickers and Van den Oever note, ‘amateur users’ provide especially rich source material for cultural historians, since their practices often reveal improvisation, failure, and appropriation.⁵⁷ Such dimensions tend to be lost in discursive accounts.

The XS4ALL homepages can make these practices visible, showing how early users experimented with forms of self-presentation, play, and community-building. The qualitative analysis that follows is illustrative rather than exhaustive. Its function is not to provide a comprehensive interpretation of the collection but to demonstrate what becomes visible once media-archaeological principles are applied to the analysis and interpretation phase of researching vernacular web pages.

Challenging Linear Narratives

Media archaeology offers a way to study media cultures by understanding past media, as Parikka notes. He writes that media cultures are best understood as ‘sedimented and layered, a fold of time and materiality where the past might be suddenly discovered anew, and the new technologies grow obsolete incredibly fast’.⁵⁸ The inherent entanglement of past and present challenges the idea that new media technologies emerge in a revolutionary manner, changing cultures and practices in their wake. This line of thinking is not unique to media archaeology; scholars across science and technology studies and media studies have long argued for moving away from linear histories of media technology that foreground innovation and rupture.⁵⁹

Building on this scholarship, Fickers and Van den Oever highlight media studies’ tendency to privilege novelty: ‘[w]hile processes of technological invention and innovation, as well as re-mediation, have been analysed in great detail, phenomena of hybridisation, habitualisation and routinization have received much less attention’.⁶⁰ Rejecting this evolutionary logic of rupture allows us to historicise the XS4ALL homepages not as the product of a revolutionary medium, but as part of a process of becoming ordinary – what Andreas Hepp refers to as *veralltäglichen*, or the becoming of everydayness.⁶¹

In this way, media archaeology’s emphasis on entanglement encourages us to view media histories as cyclical, dialectical, and renewable.⁶² Rather than searching only for new practices, we are prompted to attend to continuities and cycles, to how older cultural forms persist and are rearticulated in new environments. This approach provides us with a valuable lens for identifying and

making sense of recurring phenomena in the XS4ALL homepage collection. Many practices, for example, can be seen as continuations of longer cultural traditions: digital photo albums, often themed around family holidays or parties; guestbooks or poetry albums; digital birth announcements or Christmas cards; genealogical trees – to name a few. Such continuities do not imply that these digital practices are mere copies of analogue ones. While familiar features remained, others were adapted to suit the affordances of a website. A digital birth announcement might mirror the printed card one would expect to find in the mailbox, but the hyperlink system allowed parents to add different themed pages (a photo album, for example), continuously update information about a child's early life, or add a guestbook to arrange maternity visits. Digital photo albums, far more shareable than their analogue counterparts, were often accompanied by welcome messages to contextualise them for visitors.

Although the web was new to many of its early users, these recurring features show that people's use of it was not necessarily novel but built on existing practices and media. At the same time, they illustrate how such practices constitute what Parikka calls 'regimes of memory': structured ways of remembering and documenting, shaped by both long-standing cultural traditions and the technical affordances of the early web.⁶³ In this sense, the amateur homepage (particularly those in the everyday culture subset) functioned as a memory technology, a site where cultural memory was organised and circulated.⁶⁴

The lens of everyday culture, however, extends beyond personal memory practices. The professional subset of the collection enables the examination of small-scale online entrepreneurship, encompassing regional industries, family businesses, and hobbyists engaged in various forms of commerce. These homepages should not be read simply as novel, competent business websites; many pages blended personal and professional elements in ways that reflect a medium in transition. They often appeared playful or uncertain in purpose, with personal rhetoric, incomplete contact details, or layouts closer to business cards than to functional shops. These examples suggest that the decision to go online was tied to broader processes of normalisation, as by the mid-1990s, the growing influence of the new economy and its neoliberal ethos began to reshape local practices.⁶⁵ The coexistence of amateur memory and entrepreneurial practices within the XS4ALL collection highlights its temporal fluidity; it was simultaneously anchored in past traditions and propelled by visions of the future. Writing a linear web history would disregard this aspect.

Analysing Imaginaries and Topoi

Media-archaeological scholarship has foregrounded the importance of media imaginaries, emphasising the role of metaphors and myths, and revealing how ‘non-solid’ or ‘impossible’ media are often imagined through spectacular or fabulous language.⁶⁶ Scholarship on media imaginaries, or on the imaginary more broadly, is too vast to discuss at length in this paper. Nevertheless, we want to show how Erkki Huhtamo’s idea of media archaeology as a topos study can serve as a useful interpretative approach when examining the XS4ALL homepages.

Huhtamo adapts the rhetorical notion of topos (a commonplace or recurring theme) into a media-historical method for studying cultural clichés and recurring tropes of media use. As he writes, ‘[t]opoi should be researched as symptoms of both cultural continuities and ruptures’.⁶⁷ They emerge in specific historical contexts, are carried across media, and may take textual, visual, or design form. Topoi are not timeless archetypes but dynamic elements, reactivated and reshaped in new circumstances. This perspective resonates strongly with media archaeology’s broader critique of rupture-centred narratives; rather than treating metaphors or tropes as surface embellishments, they are seen as active frameworks through which users made sense of the web.

At the micro level, we can recognise topoi in the XS4ALL collection as motifs that circulated widely in 1990s web culture: under-construction GIFs, welcome banners, hit counters, or blinking guestbook icons. These devices signalled belonging to a shared repertoire of web vernaculars, what Olia Lialina later described as the vernacular web.⁶⁸ At the macro level, broader imaginaries framed what the web was and could be. In the Dutch context, recurring metaphors included *de digitale snelweg* (the information superhighway) and the digital city.⁶⁹ These imaginaries were not created by users alone but were part of a wider socio-technical discourse that shaped both policy and design. Crucially, the homepage itself can be seen as a topos or an imaginary. Its very naming as a *home* framed the web as a space of selfhood, identity, and presence online. Mallapragada has shown how the homepage foregrounds affective and personal notions, offering users a space of their own on the network.⁷⁰

This metaphorical logic shaped how users imagined their relation to the web and informed the design of certain pages. For example, some pages replicate a literal home through which one could ‘walk’ by clicking hyperlinks to different rooms – a design practice that echoed similar navigational ecologies in earlier BBS environments.⁷¹ In this sense, the homepage embodies the entanglement of

imaginaries and topoi; a recurring cultural script that at once drew on older associations of domesticity, authorship, and ownership, and gave shape to new practices of self-presentation online.

At both levels, the aim is not to catalogue surface aesthetics but to show how recurring tropes and metaphors structured historical experience. By tracing everyday practices alongside topoi and imaginaries, we can write a cultural history of the web that is non-linear, grounded in use, and attentive to the continuities between past and present media cultures.

Conclusion and Discussion

Through three media-archaeological principles, this article has traced the epistemic processes at work when researching an archived web collection and has shown how these processes shape each stage of writing a cultural media history. The XS4ALL homepages constitute a distinctive case, both in terms of content and archival context, yet the dynamics analysed here resonate with broader methodological debates previously articulated within web studies. The interpretation of modalities and layers, understood as part of the materiality of the archived web, has been examined by scholars such as Ankerson and Jansma, both of whom foreground software and interfaces as historically meaningful artefacts and sources.⁷² The technical and human dynamics shaping the archival and heritage-making processes of websites, as well as the plurality of actors involved, have been addressed more broadly by Schafer et al. albeit primarily through the lens of science and technology studies.⁷³ Questions of temporal and spatial inconsistency are well-established within web studies scholarship and have been analysed through frameworks of source criticism and archival negotiation by Brügger, alongside methodological reflections on the preparatory interpretive and technical labour involved in web archive research by Gebeil and Thièvre.⁷⁴ Taken together, the existing body of work underscores the extent of source criticism, triangulation, iterative experimentation, and often collaborative inquiry demanded by web-historical research, and thus the need for a certain level of digital archive literacy – or web archive literacy – to engage with these sources.

The present article does not claim the themes above as new but contributes to this body of scholarship in two ways. First, it offers a situated methodological reflection grounded in the Dutch context, working within the institutional and infrastructural conditions of the KB and through a single, historically anchored collection approached via a specific actor. Second, on a broader level, it

shows how selected principles and perspectives from media archaeology can be adapted, not as abstract theory but as practical guidance for historical research, proving particularly apt for the study of past vernacular web practices. In this respect, the article aligns most closely with empirically grounded studies such as those by Jansma, as well as Paßmann and Gerzen, which approach archived web artefacts as traces of practice.⁷⁵ It extends this line of work by integrating media-archaeological reflection directly into the act of writing web history.

Our analysis clarifies that applying media-archaeological principles to archived web collections not only structures historical inquiry but also foregrounds questions of situatedness, power, and subjectivity. Skågeby and Rahm extend the notion of medium-specificity by embedding it within intersectional and ethical concerns, asking whom the medium serves or excludes, and which histories or practices remain hidden.⁷⁶ Applied to XS4ALL, such questions foreground who the ‘all’ in XS4ALL truly encompasses. Despite its branding as a people’s network, the community was dominated by young, highly educated men, often bound by insider humour that excluded others.⁷⁷ Its reputation as an elite provider was reinforced by its appeal to heavy users and early adopters, a factor that contributed to its later acquisition by KPN.⁷⁸ Curatorial choices further shaped what has survived, privileging homepages that were less fragmented and easier to capture. Yet, as Fickers and Van den Oever remind us, moments of failure or disappearance are equally revealing from a media-archaeological perspective.⁷⁹ Tools and infrastructures add another excavated layer here; designed for more modern web architectures, they risk producing further ‘data loss’, highlighting the need for transparency about preprocessing practices in archived web data.⁸⁰

A second point concerns historical canonisation. While UNESCO has recognised XS4ALL as part of Dutch digital heritage, the provider has also actively cultivated its own legacy as a pioneer.⁸¹ This canonisation reinforces a distinctly Amsterdam-centred narrative of the internet, foregrounding the urban, activist origins of XS4ALL while sidelining other regional or more mundane histories, including ISPs founded before XS4ALL.⁸² What counts as heritage, then, is never self-evident but the outcome of situated processes of valuation within social and institutional logics.

Finally, historical re-enactment – central to the experimental media-archaeological method – proved only partially possible in this study. Archived websites could be replayed, but only on contemporary screens and software. Obsolete browsers, software plug-ins, and slower loading times – the very conditions that shaped the early web experience – could not be reproduced at this point at the KB. This tacit or sensorial knowledge thus fell outside the scope of analysis. Recognising these

limitations alongside the other challenges of working with archived web data, such as source criticism and chronology, raises questions of authenticity and accountability. However, we argue that partiality is not a failure but the condition of retroactive web archives, or archives collected in a state of urgency – what Ian Milligan has theorised along the lines of crisis collecting.⁸³ These sources remain invaluable precisely because they preserve traces that would otherwise disappear. The XS4ALL homepages exist because of such selective and subjective interventions. To approach them through a media-archaeological lens is not to overcome these conditions but to excavate them, acknowledging their limits while affirming their indispensability to writing cultural histories of the web.

Acknowledgements

The authors thank Prof. Dr. Susan Aasman (University of Groningen, Centre for Media and Journalism Studies) for her valuable feedback on an earlier version of the manuscript. They also thank the reviewers and editors for their constructive comments and pointed suggestions, which helped strengthen the article considerably. The authors are grateful to the staff of the Koninklijke Bibliotheek (KB) for their assistance during the research process, including colleagues from the IT, legal, and collection teams.

Declaration of conflicting interest

The authors declare there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding statement

The authors received no external financial support for the research, authorship, and/or publication of this article. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethical approval and informed consent statement

This research project was reviewed and approved by the Research Ethics Committee (CETO) of the Faculty of Arts, University of Groningen. The Committee confirmed that the research protocol complies with internationally recognised standards.

Data availability statement

This study draws on archived websites from the XS4ALL homepage collection preserved by the Koninklijke Bibliotheek (KB), the National Library of the Netherlands. Research access was granted in accordance with the KB's guidelines and usage policies. The data is partially publicly available due to copyright and privacy restrictions, but may be consulted under the same conditions outlined in this article.

Notes

1. Jussi Parikka, *What is Media Archaeology?* (Polity Press, 2013), 167.
2. Ian Milligan, *History in the Age of Abundance? How the Web is Transforming Historical Research* (McGill-Queen's University Press, 2019).
3. Jesper Verhoef, "A 1980s and 1990s Media History Manifesto," *TMG Journal for Media History* 26, no. 2 (2023): 1–34.
4. Eveline Vlassenroot et al., "The Evolving Landscape of Web and Social Media Archiving: A Comprehensive Review of Current Practices in National Libraries and Archives," *International Journal of Digital Humanities* (2025): 1–35.
5. "The Datafied Web – RESAW Conference 2025," Universität Siegen, accessed September 11, 2025, <https://www.uni-siegen.de/en/2025/resaw-2025-the-datafied-web>.
6. Megan Sapnar Ankerson, "Read/Write the Digital Archive: Strategies for Historical Web Research," in *Digital Research Confidential: The Secrets of Studying Behavior Online*, ed. Eszter Hargittai and Christian Sandvig (MIT Press, 2015), 29–54; Niels Brügger, *The Archived Web: Doing History in a Digital Age* (MIT Press, 2018), 119–48; Valérie Schafer et al., "Negotiating the Web of the Past: Web Archiving, Governance

- and STS,” *French Journal for Media Research*, no. 6 (2016): 1–23; Johannes Paßmann and Lisa Gerzen, “Follow the Updates! Reconstructing Past Practices with Web Archive Data,” *Internet Histories* 8, no. 3 (2024): 213–28; Sophie Gebeil and Jérôme Thièvre, “Screens in Struggle: From Archived Web Corpus to Readable Data for History Research,” in *The Routledge Companion to Transnational Web Archive Studies*, ed. Susan Aasman et al. (Routledge, 2025), 361–78.
7. “Betere regelgeving voor behoud digitaal erfgoed,” Rijksoverheid, accessed January 14, 2026, <https://www.rijksoverheid.nl/actueel/nieuws/2025/10/02/betere-regelgeving-voor-behoud-digitaal-erfgoed>.
 8. For scholarship on Dutch web history, see Esther Weltevrede and Anne Helmond, “Where do Bloggers Blog? Platform Transitions within the Historical Dutch Blogosphere,” *First Monday* (2012); Caroline Nevejan and Alexander Badenoch, “How Amsterdam Invented the Internet: European Networks of Significance, 1980–1995,” in *Hacking Europe: From Computer Cultures to Demoscenes*, ed. Gerard Alberts and Ruth Oldenziel (Springer, 2014), 189–217; Nathalie Fridzema et al., “A Decade of Transformation Discourse: Sociotechnical Imaginaries of the Dutch Web Between 1994–2004,” in *Exploring the Archived Web During a Highly Transformative Age: Proceedings of the 5th international RESAW conference, Marseille, June 2023*, ed. Sophie Gebeil and Jean-Christophe Peyssard (Firenze University Press, 2024), 141–61. For a methodological approach of working with Dutch web collections, see Jesper Verhoef, “Doing LGBTQ Internet Histories Justice: A Queer Web Archive Manifesto,” *Internet Histories*, 9, nos. 1–2 (2025): 180–97.
 9. Susan Aasman et al., “Web Archaeology: An Introduction,” *TMG Journal for Media History* 22, no. 1 (2019): 1–5; Robert Jansma, “Metadata Dating the Digital City: A Software Archaeological Approach,” *Internet Histories* 6, no. 4 (2022): 452–72; Kees Teszelszky, “Web Archaeology in The Netherlands: The Selection and Harvest of the Dutch Web Incunables of Provider Euronet (1994–2000),” *Internet Histories* 3, no. 2 (2019): 180–94.
 10. “Webcollectie XS4ALL-homepages (1993–2001),” KB, accessed September 11, 2025, <https://www.kb.nl/over-ons/projecten/xs4all-homepages-archivering>.
 11. “Hack-Tic Magazine Archive (1989–1994),” Hack-Tic, accessed September 9, 2025, <https://www.hacktic.nl/>.
 12. Nathalie Fridzema et al., “Creating the Early Dutch Web: XS4ALL’s Curatorial Role as Pioneer Community,” in *Silicon Dawn: Creative Computing in Europe 1970–2000*, ed. Alex Wade et al. (University of Amsterdam Press, forthcoming).
 13. “De ‘webwiegendrukken’ van het Nederlandse internet: webcollectie XS4ALL-homepages,” UNESCO, accessed September 11, 2025, <https://www.unesco.nl/nederlands-memory-of-the-world-register/de-webwiegendrukken-van-het-nederlandse-internet>.

14. Ankerson, "Read/Write the Digital Archive," 29–54; Gebeil and Thièvre, "Screens in Struggle," 361–78.
15. Parikka, *What is Media Archaeology?*, 163.
16. Parikka, *What is Media Archaeology?*, 164.
17. Parikka, *What is Media Archaeology?*, 87. Our emphasis.
18. Brügger, *The Archived Web*, 119–48.
19. Niels Brügger, "Web History and the Web as a Historical Source," *Zeithistorische Forschungen/Studies in Contemporary History* 9, no. 2 (2012): 316–25.
20. Niels Brügger, "The Archived Website and Website Philology: A New Type of Historical Document?," *Nordicom Review* 29, no. 2 (2008): 155–75.
21. Parikka, *What is Media Archaeology?*, 15.
22. Parikka, *What is Media Archaeology?*, 87.
23. Wolfgang Ernst, "Media Archaeography: Method and Machine versus History and Narrative of Media," in *Media Archaeology: Approaches, Applications, and Implications*, ed. Erkki Huhtamo and Jussi Parikka (University of California Press, 2011), 242.
24. "Voorbeelden van onze Speciale Webcollecties," KB, accessed September 11, 2025, <https://www.kb.nl/onderzoeken-vinden/bijzondere-collecties/webcollectie/voorbeelden-van-onze-speciale-webcollecties>.
25. Fridzema et al., "Creating the Early Dutch Web."
26. "XS4ALL Home Pages (1993–2001) Web Collection," KB, accessed September 11, 2025, <https://www.kb.nl/en/over-ons/projecten/xs4all-homepages-archivering>.
27. "Legal Aspects in Web Archiving," KB, accessed September 11, 2025, <https://www.kb.nl/en/over-ons/expertises/webarchivering/juridische-aspecten>.
28. Iris Geldermans, "No Longer XS4ALL," KB Lab, July 25, 2024, <https://lab.kb.nl/about-us/blog/no-longer-xs4all>.
29. Peter de Bode and Kees Teszelsky, *Collectiebeschrijving XS4ALL-homepages* (KB - Nationale Bibliotheek, 2021).
30. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*, 21.
31. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*, 8.
32. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*, 25.
33. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*, 69.
34. "Overview," Web Curator Tool, accessed September 11, 2025, <https://webcuratortool.org/>.
35. Brügger, *The Archived Web*, 128.

36. "Using the Wayback Machine," Internet Archive, accessed September 11, 2025, <https://help.archive.org/help/using-the-wayback-machine/>.
37. Gebeil and Thièvre, "Screens in Struggle," 364.
38. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*.
39. Megan Sapnar Ankerson, *Dot-Com Design: The Rise of a Usable, Social, Commercial Web* (New York University Press, 2018).
40. Brianna Kennedy and Robert Thornberg, "Deduction, Induction, and Abduction," in *The SAGE Handbook of Qualitative Data Collection*, ed. Uwe Flick (SAGE Publications, 2017).
41. "Webcollectie XS4ALL-homepages (1993-2001)," KB, accessed September 11, 2025, <https://www.kb.nl/over-ons/projecten/xs4all-homepages-archivering..>
42. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-editedhomepages*.
43. Zack Lischer-Katz, "Studying the Materiality of Media out Archives in the Age of Digitization: Forensics, Infrastructures and Ecologies," *First Monday* 22, no. 1 (2017).
44. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 14.
45. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 28.
46. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 37.
47. "About/Visit," Media Archaeology Lab, accessed September 11, 2025, <https://www.mediaarchaeologylab.com/about>.
48. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 81.
49. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 28.
50. "KB_datalab," KB, accessed September 11, 2025, <https://www.kb.nl/over-ons/diensten/kb-datalab>.
51. Anne Helmond, "A Historiography of the Hyperlink: Periodizing the Web Through the Changing Role of the Hyperlink," in *The SAGE Handbook of Web History*, ed. Niels Brügger and Ian Milligan (SAGE, 2019), 153–67.
52. "SolrWayback read me," Netarchivesuite Github, accessed September 11, 2025, <https://github.com/netarchivesuite/solrwayback>.
53. Anchor hyperlinks navigate users to other sections within the same webpage.
54. Iris Geldermans, "Let's Get Some Data - Link Analysis Part 1," *KB Lab*, April 15, 2022, <https://lab.kb.nl/about-us/blog/lets-get-some-data-link-analysis-part-1>.
55. De Bode and Teszelsky, *Collectiebeschrijving XS4ALL-homepages*, 44.
56. Wanda Strauven, "Media archaeology as Laboratory for History Writing and Theory Making," in *New Media Archaeologies*, ed. Ben Roberts and Mark Goodall (Amsterdam University Press, 2019), 28.

57. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 33, 37.
58. Parikka, *What is Media Archaeology?*, 3.
59. Carolyn Marvin, *When Old Technologies Were New: Thinking about Electric Communication in the Late Nineteenth Century* (Oxford University Press, 1988); Brian Winston, *Media, Technology and Society: A History: From the Telegraph to the Internet* (Routledge, 2002); Wendy Hui Kyong Chun et al., eds., *New Media, Old Media: A History and Theory Reader* (Routledge, 2004); Lisa Gitelman, *Always Already New: Media, History, and the Data of Culture* (MIT Press, 2008).
60. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 53.
61. Andreas Hepp, "Curators of Digital Futures: The Life Cycle of Pioneer Communities," *New Media & Society* 27, no. 9 (2024): 5390–409.
62. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*.
63. Parikka, *What is Media Archaeology?*, 2.
64. Nancy van House and Elizabeth F. Churchill, "Technologies of Memory: Key Issues and Critical perspectives," *Memory Studies* 1, no. 3 (2008): 295–310.
65. Nathalie Fridzema et al., "Small-scale Entrepreneurship on the Early Web: Socio-Economical Practices of Local/Regional Businesses," *AoIR Selected Papers of Internet Research AOIR2025* (2025).
66. The term 'non-solid' is used by Parikka, *What is Media Archaeology?*, 62; The term 'impossible' is used by Eric Kluitenberg, "On the Archaeology of Imaginary Media," in *Media Archaeology: Approaches, Applications, and Implications*, ed. Erkki Huhtamo and Jussi Parikka (University of California Press, 2011), 48.
67. Huhtamo, "Dismantling the Fairy Engine," 34.
68. Olia Lialina and Dragan Espenschied, eds., *Digital Folklore Reader* (merz & solitude, 2009).
69. Fridzema et al., "A Decade of Transformation Discourse," 141–61.
70. Madhavi Mallapragada, "Cultural Historiography of the 'Homepage'," in *The SAGE Handbook of Web History*, ed. Niels Brügger and Ian Millian (SAGE Publications, 2019).
71. Kevin Driscoll, *The Modem World: A Prehistory of Social Media* (Yale University Press, 2022).
72. Ankerson, "Read/Write the Digital Archive," 29–54; Jansma, "Metadata Dating the Digital City," 452–72.
73. Schafer et al., "Negotiating the Web of the Past," 1–23.
74. Brügger, *The Archived Web*, 119–48; Gebeil and Thièvre, "Screens in Struggle," 361–78.
75. Jansma, "Metadata Dating the Digital City," 452–72; Paßmann and Gerzen, "Follow the Updates!," 213–28.
76. Jörgen Skågeby and Lina Rahm, "What is Feminist Media Archaeology?" *Communication+ 1* 7, no. 1 (2018).

77. Caroline Nevejan and Alexander Badenoch, “How Amsterdam Invented the Internet: European Networks of Significance, 1980–1995,” in *Hacking Europe: From Computer Cultures to Demoscenes*, ed. Gerard Alberts and Ruth Oldenziel (Springer, 2014); Maarten Reijnders, *De hackers die Nederland veranderden. De spannende geschiedenis van XS4ALL* (Nieuw Amsterdam, 2023); Fridzema et al., “Creating the Early Dutch Web.”
78. Kemenade, “KPN en ING doorzetter: Ron van Kemenade.”
79. Fickers and Van den Oever, *Doing Experimental Media Archaeology: Theory*, 37.
80. “Data loss (DALOSS): The Politics of Disappearance, Destruction and Dispossession in Digital Societies,” University of Copenhagen, accessed September 11, 2025, <https://artsandculturalstudies.ku.dk/research/daloss/>.
81. Fridzema et al., “Creating the Early Dutch Web.”
82. Reijnders, *De hackers die Nederland veranderden*.
83. Ian Milligan, *Averting the Digital Dark Age: How Archivists, Librarians, and Technologists Built the Web a Memory* (Johns Hopkins University Press, 2024).

Biographies

Nathalie Fridzema is a PhD candidate at the Centre for Media and Journalism Studies, University of Groningen. Her dissertation examines the cultural history of the early Dutch web (1994–2004), focusing on web imaginaries, everyday practices, and regional histories often overlooked in dominant accounts of internet history. Combining qualitative and quantitative approaches, she draws on media archaeology, web anthropology, and archival studies to explore how the early web has been imagined, curated, and preserved. She is particularly interested in the vernacular web and in how local and amateur practices shape digital heritage collections.

University of Groningen, Centre for Media and Journalism Studies

n.fridzema@rug.nl

31 6 57005118

ORCID: 0009-0006-3044-8246

Iris Geldermans is a researcher in web archiving and finds herself at the heart of web archiving at the KB | National Library of the Netherlands. At the crossroads of technique and collection, she aids colleagues and researchers with a wide variety of questions. Whether it is analysing and researching

archived data or finding ways to archive websites by tinkering with harvest profiles, she loves solving the new puzzles that the web archiving field brings to her every day.

KB | National Library of the Netherlands

iris.geldermans@kb.nl

31 6 39645232

<https://lab.kb.nl/person/iris-geldermans>

TMG Journal for Media History

Volume 29 No (1)/2026

DOI

<https://doi.org/10.18146/tmg.944>

PUBLISHER

Netherlands Institute for Sound & Vision

COPYRIGHT

Each article is copyrighted © by its author(s) and is published under license from the author(s). When a paper is accepted for publication, authors will be requested to agree with the Creative Commons Attribution-ShareAlike 4.0 International License.