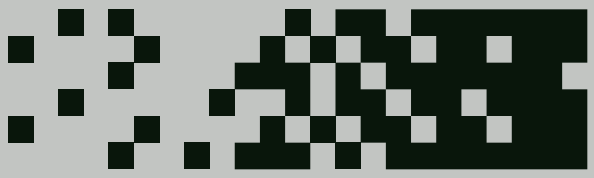
Vanishing Culture

SAVE

Vanishing Culture:

A Report on Our Fragile Cultural Record

edited by



Luca Messarra[1]
Chris Freeland[2]
and Juliya Ziskina[3]

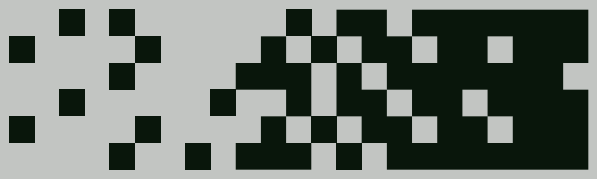
1 ▪ Luca Messarra (messarra@stanford.edu) is a Public Humanities Fellow at the Internet Archive and an English PhD candidate at Stanford University, specializing in contemporary American literature, the history of text technologies, and the digital humanities. His dissertation explores how print on demand has revolutionized literary publishing and experimentation. His creative scholarship may be found under his imprint, Undocumented Press.

2 ▪ Chris Freeland (chrisfreeland@archive.org) is the Director of Library Services at the Internet Archive.

3 ▪ Juliya Ziskina (juliya@archive.org) is a policy attorney at the Internet Archive and the eBook Study Group.

Contents

Part 1:



Media Preservation and the Production of Public Memory

Luca Messarra

Public Humanities Fellow,
Internet Archive

“Look at this [pocketwatch]. It’s worthless. Ten dollars from a vendor in the street. But I take it, I bury it in the sand for a thousand years, it becomes priceless.”

- René Belloq, *Raiders of the Lost Ark*[4]

“There is no digital equivalent to that decades-old pile of *Life* or *National Geographic* magazines in the basement or attic. Changes in computing technology will ensure that over relatively short periods of time, both the media and the technical format of old digital materials will become unusable. Keeping digital resources for use by future generations will require conscious effort and continual investment.”

- Dale Flecker, *Harvard University Library*[5]

4 ▪ Steven Spielberg, dir. *Raiders of the Lost Ark*. Paramount Pictures, 1981.

5 ▪ Dale Flecker, “Preserving Digital Periodicals,” in *Appendices: Preserving Our Digital Heritage: Plan for the National Digital Information Infrastructure and Preservation Program* (Library of Congress, 2002), 26.

I. The Digital Memory Hole

The long history of hardware, software, and media obsolescence, from the papyrus scroll to the Commodore 64, tells us that if archivists and librarians do not actively collect and preserve cultural materials, then they risk being forever lost in the sands of time. Unlike Belloq's hypothetical watch in the sand, digital culture[6] is valuable *now*, and is vanishing from public access at an alarming rate: social media posts are deleted by censors and malicious agents seeking to erase political history;[7] culturally significant journalism such as MTV News,[8] alongside music,[9] films,[10] and television shows,[11] are abruptly taken offline by media conglomerates without explanation; and landmark video games created for old hardware fade from our memory under restrictive and lengthy copyright terms.[12] Furthermore, analog materials such as VHS tapes, 78rpm recordings, and filmstrips are deteriorating moment by moment, and require urgent attention to ensure their survival.

"Historical oblivion is the default, not the exception" to the human record, writes game designer Jordan Mechner in his contribution to this report.[13] Be it natural elements like fire or water, negligent or intentional people, or simple forgetfulness, practically all human expression will disappear or change without human intervention. Only through acts of repair and digitization will materials such as a grandmother's cookbook, a groundbreaking game ahead of its time, or endangered languages continue to survive, and only through access will their survival prove meaningful to the next generation. In response to the risk of materials falling into oblivion, humanity has built libraries and archives big and small, and developed rigorous preservation practices, so that the past need not be so easily lost. These institutions have not always collected fairly, nor have they always been accessible.[14] Yet now more than ever, with the help of dedicated individuals and communities, these institutions seek to ethically collect and preserve culture broadly. They also offer digital lending programs, so that anyone in the world with an internet connection can enjoy and learn from the past, not just those with the means to visit a physical location.

6 ▪ "Culture" is broadly defined in this report as any publicly disseminated human creation.

7 ▪ See e.g. PolitiTweet for an archive of deleted tweets by public figures and organizations. The Library of Congress previously archived all public tweets from 2006-2017, but has since transitioned to a "very selective" collecting protocol. See Laurel Wamsley, "Library Of Congress Will No Longer Archive Every Tweet," *NPR*, December 26, 2017.

8 ▪ Todd Sprangler, "MTV News Website Goes Dark, Archives Pulled Offline," *Variety*, June 24, 2024.

9 ▪ Ari Herstand, "Takedown Fails: Artists Are Seeing Their Music Removed From DSPs for Streaming Fraud They Didn't Commit," *Variety*, April 9, 2024.

10 ▪ Zach Schonfeld, "Beware Hollywood's digital demolition: it's as if your favourite films and TV shows never existed," *The Guardian*, October 1, 2024.

11 ▪ Brad Adgate, "Warner Bros. Discovery Shuts Down Cartoon Network Website," *Forbes*, August 12, 2024.

12 ▪ *Techdirt* frequently publishes on copyright inhibiting video game reissuing. See e.g. Timothy Geigner, "Copyright Self Censorship Denies Us Another Updated Version Of An Abandoned Game," *Techdirt*, July 27, 2023.

13 ▪ Jordan Mechner, "Preserving Gaming History," in *Vanishing Culture: A Report on Our Fragile Cultural Record*, ed. Luca Messarra, Chris Freeland, and Juliya Ziskina (Internet Archive, 2024), 100.

14 ▪ See Andrew Pettegree and Arthur der Weduwen, *The Library: A Fragile History* (Basic Books, 2021) for an in-depth history of libraries.

But as rights holders increasingly opt for provisional access by way of streaming platforms and temporary licenses, culture is vanishing from public access without any indication that it will return. These distribution methods inhibit the critical ability of individuals and institutions to own, and therefore preserve their historical record, placing public memory in the hands of media conglomerates and their interests. The absence of ownership not only erodes the public's ability to enjoy, learn from, and collaboratively build their own diverse cultural history: it creates an information ecosystem where history can be seamlessly rewritten by revising and deleting public information and popular content.

How Did We Get Here 

Corporate interests interfering with the library mission—that is, preserving and making culture accessible to the public—is far from new. Historically, legal protections for crucial library services such as interlibrary loan, electronic reserves for students, photocopying, and reformatting exceptions for patrons with print disabilities, were only secured after publishers attempted (and failed) to challenge these practices in court.[15] As copyright experts Kyle Courtney and Juliya Ziskina wrote in 2023: "The historical conflict between libraries and publishers reveals a predictable pattern. Libraries are fast to adopt new ways of providing greater access to knowledge to their patrons. Publishers react by obstructing libraries' efforts."[16]

Public-serving libraries and archives in the United States came of age and acquired their legal protections over the nineteenth and twentieth centuries, a media environment dominated by print books. The book is a marvel of human engineering: it is sturdy, portable, and easy to mass reproduce; it remains among the best long term data storage media ever created. We have bound works of ink on parchment that are over a thousand years old, whose data remains nearly as parsable as the day it was first inscribed. Left alone in a temperature-controlled room, a book made today will remain usable for hundreds of years.[17] Cumulatively, these assumptions about book longevity and distribution have informed the way the public and lawmakers generally think about the work that libraries do: putting books on shelves, lending them out, and applying archival adhesive and book tape as needed.

15 ▪ For a systematic overview of these obstructions, see Kyle Courtney and Juliya Ziskina, "The Publisher Playbook: A Brief History of the Publishing Industry's Obstruction of the Library Mission," *HCA Scholarly Articles* (2023). Pre-print.

16 ▪ Courtney and Ziskina, "The Publisher Playbook," 21.

17 ▪ "The Deterioration and Preservation of Paper: Some Essential Facts," *Library of Congress*.

But we no longer live in a print-dominant world. Libraries and archives collect and preserve television broadcasts, films, oral history recordings, ebooks, web pages, social media posts, and more. Yet, anyone who has ever tried to use a floppy disk lately—or who has discovered that their favorite meme has disappeared from Tumblr or TikTok (or that their favorite platform has all together shut down like Vine in 2017)—knows that digital materials tend to become obsolete, disappear, and physically decay far faster than books on bookshelves.[18] As Dale Flecker, former Associate Director for Planning and Systems at Harvard University Library describes, the constantly evolving nature of software and hardware creates a media environment in perpetual need of "conscious effort and continual investment" so as to remain publicly accessible. That effort and investment is the sort of labor-intensive and unprofitable work that caring individuals and public-serving institutions have long demonstrated a willingness to do, even when there is no clear economic gain in the act of preservation.

18 ▪ For more on Vine, see Casey Newton, "Why Vine Died," *The Verge*, October 28, 2016. Notably, Twitter had created a Vine archive in 2017 so that users could browse old content. That archive was taken down in 2019.

> **"I left YouTube for a while in 2022, when Scholastic, one of the largest children's book publishers on earth, tried to get my channel deleted. Turns out they bought the assets of a defunct filmstrip publisher whose work I was trying to save. So not only had no one preserved these things, but a corporation hoarding bankruptcy assets now threatened the very point of preservation in the first place: making history available for viewing."**
>
> **- Mark O'Brien,** "On Filmstrips"

As the next section of this report will show, publishers across all media have demonstrated time after time that they will not preserve or make their publications commercially available when there is no economic incentive to do so. Worse, they will use the law to shut down non-commercial preservation and accessibility efforts that seek to extend the life of cultural materials.[19] Those activities, unfortunately, make perfect sense for businesses beholden to shareholder interests. After all, the fundamental principle of a large publisher seeking profit and growth is to keep costs as low as possible while keeping sales as high as possible. It is entirely logical for publishers to abandon publications when the economic incentive to continue distributing and preserving them is gone. Further, it

19 ▪ For discussions of recent lawsuits against the Internet Archive, see Maria Bustillos, "Publishers Are Taking the Internet to Court," *The Nation*, September 10, 2020 and Kate Knibbs, "The Internet Archive's Fight to Save Itself," *Wired*, September 27, 2024.

makes business sense for rights holders to retain their long copyrights—now lasting the life of the author plus seventy years, or ninety-five years in the case of works made for hire—in the event that a reissue may one day turn out profitable.[20]

Preserving and making works accessible within a profit-motivated distribution paradigm was manageable for libraries as long as human expression was predominantly published in book form. Over the twentieth century, the law empowered libraries to purchase, lend, and repair books so that they could be made available long after publishers decided to stop distributing them. These laws created a shaky balance between corporate and public interests in cultural preservation: publishers could cease book production whenever they saw fit, and libraries would do the work of collecting and preserving books for future generations, allowing the public to reevaluate the continued preservation of these works in a different historical moment. This balance, predicated upon media ownership, not only benefited the public by giving them access to old works, it also relieved publishers from the burden of keeping authors' works in-print, allowing publishers to focus on bringing new books to life, and giving less commercially successful authors the chance to continue being read.[21]

Opportunities for ownership across all media rapidly decreased over the course of the 2010s as digital content increasingly became distributed through streaming platforms alone, and as publishers' moved toward subscription-based licensing deals for libraries.[22] These changes in distribution have heavily shifted the balance of cultural posterity in favor of publishers. Juliya Ziskina details the crux of this issue, particularly in the case of ebook licensing, as such:

> The shift from ownership toward licensing opened the door to the substitution of statutory property rights (such as the right to acquire, use, and dispose of property), replacing them with unilateral contract terms. As a result of this shift, publishers and platforms now routinely attempt to assert control over almost all library activities related to ebooks, including how, where, when, and to whom they can be shared.[23]

Instead of owning copies of digital materials, individuals and public-serving institutions now pay for temporary and conditional access to digital content. Even when consumers hit 'buy' on their favorite platform, what they are most often buying is a limited right to stream that content—using only the publisher's software of choice—with the caveat that the content can be removed or changed at any time.[24] While streaming may

20 ▪ For a robust discussion of how copyright facilitates cultural vanishing, see Mark A. Lemley, "Disappearing Content," *Boston University Law Review* 101, no. 4 (2021).

21 ▪ Despite the benefits given to authors by this balance, publishers have routinely attempted to control what readers and institutions do with publications after first sale. See Sarah Lamdan et al., "From Physical Book Sales to Ebook Licensing – An Opportunity for Publishers to Achieve Long-Held Goals" in *The Anti-Ownership Ebook Economy: How Publishers and Platforms Have Reshaped the Way We Read in the Digital Age* (Engelberg Center on Innovation Law & Policy, 2023) for the long legal history of the first sale doctrine, which codified the right for libraries to lend books.

22 ▪ Libraries used to be able to buy perpetual licenses for ebooks from the Big Five publishers, but by 2020, all five of these publishers moved over to limited licenses. Sarah Lamdan et al., *The Anti-Ownership Ebook Economy*, 31.

23 ▪ Juliya Ziskina, *Toward a New Access Paradigm: Digital Ownership for Libraries and the Public* (Library Futures, 2023), 2.

24 ▪ See Aaron Perzanowski and Jason Schultz, *The End of Ownership: Personal Property in the Digital Economy* (MIT Press, 2016), particularly chapter five, "The 'Buy Now' Lie," for a rigorous critique of the lack of ownership opportunities in today's digital economy.

offer convenience for consumers in the present who do not want to buy physical media, it poses much larger problems for the sustainability and possibility of a cultural record built by public interests.[25]

To summarize a few of these problems, specifically around ebooks:[26] proprietary file formats and DRM, coupled with the Big Five publishing houses' partnerships with Apple and Amazon, create ebook monopolies and monopsonies, locking readers into viewing books on their platforms, and preventing readers from engaging in the time-honored tradition of casually lending a good book to a good friend;[27] those platforms pose data privacy issues to consumers by forcing them to view books using software that mines and potentially sells their data.[28] Additionally, publishers offer temporary licenses to libraries (which often cost more than the physical books themselves), forcing libraries with limited funds to make difficult decisions about what content they are willing to acquire and for how long;[29] and in peak Orwellian fashion, publishers have retroactively applied changes to and deleted already-purchased ebooks.[30]

25 ▪ The ramifications of the licensing problem are too vast and complex to be discussed here. Fortunately, the literature on licensing's effects on individual consumers, creators, and public-serving institutions is equally extensive. From the point of view of consumers, see Perzanowski and Schultz, *The End of Ownership: Personal Property in the Digital Economy*. For an emphasis on creators and more, see Glyn Moody, *Walled Culture: How Big Content Uses Technology and the Law to Lock Down Culture and Keep Creators Poor* (BTF Press, 2022). For a technical overview from the point of view of libraries, including details about the different licensing agreements the Big Five book publishers offer to libraries, see Rachel Noorda and Kathi Inman Berens, *Digital Public Library Ecosystem 2023* (ALA, 2023). Lastly, Lila Bailey and Michael Menna have surveyed librarians' perspectives on these issues in *Securing Digital Rights For Libraries: Towards an Affirmative Policy Agenda for a Better Internet* (Internet Archive, 2022).

26 ▪ Ebooks are centered in this discussion in part because librarians and their advocates have made significant efforts to make transparent and critique this particular transformation in book publishing and access. More research is needed on sound and image distribution, but these industries are notoriously more difficult to research because of corporate obstruction. See e.g. Meredith Rose, *Streaming in the Dark: Where Music Listeners' Money Goes—and Doesn't* (Public Knowledge, 2023) for a particular emphasis on the "information asymmetries" within the music streaming industry that inhibit research and disadvantage artists and consumers.

27 ▪ See "Digital Rights Management Technology Creates Lock-In" in Lamdan et al., *The Anti-Ownership Ebook Economy*.

28 ▪ For a discussion of Amazon's data tracking of Kindle users, see Kari Paul, "'They know us better than we know ourselves': how Amazon tracked my last two years of reading" *The Guardian*, February 3, 2020. See also Library Freedom's "Vendor Privacy Scorecard" for an evaluation of major ebook vendors' data privacy practices. Lastly, see Defective by Design's "DRM Frequently Asked Questions" for an introduction to DRM.

29 ▪ See e.g. Susan Haigh, "Libraries struggle to afford the demand for e-books and seek new state laws in fight with publishers," *Associated Press*, last updated March 12, 2024.

30 ▪ Roald Dahl, R.L. Stine, and Agatha Christie are just some of the authors who have had their works retroactively changed after being 'purchased' by consumers. See Reggie Ugwu, "It's Their Content, You're Just Licensing it," *The New York Times*, April 4, 2023.

"In many instances, these important auxiliary texts are out of print, which means access via the Internet Archive is the best way scholars not located in the West might ever be able to access them [...] With the recent ruling in the publishers' lawsuit, I fear researchers, journalists, writers and other people on or from the African continent who investigate and curate knowledge for the public have lost a valuable tool for countering false narratives."

- Helen Nde (interviewed by Caralee Adams),
"Preserving African Folktales: Interview with Laura Gibbs and Helen Nde"

Perhaps above all is a concern for an equitable historical record: when the sole means of accessing culture is through licensing, we leave our cultural memory at the whim of corporate and shareholder interests, who have not hesitated to ax creators and communities from the record when they see no financial gain in their continued availability. Today's distribution practices fail the public, from which the title of "publisher" derives its name. The word "publish" comes from the Latin "publicare," meaning "to show or tell to the people, to impart to the public, make public or common."[31] In the absence of sustained public access, *public*-ation is a misnomer for works distributed in the platform era. What were once publishers might now be more appropriately called data brokers, whose interest is in maximizing profit by limiting public access and collecting user data.[32]

Indeed, the profitable datafication of cultural consumption has already taken over the connected television (CTV) industry (e.g. smart TVs like the Roku TV and its supported streaming platforms) at the expense of individual privacy. In an October 2024 report published by the Center for Digital Democracy titled *How TV Watches Us: Commercial Surveillance in the Streaming Era*, Jeff Chester, executive director at the Center for Digital Democracy, and Kathryn C. Montgomery, Professor Emerita at American University's School of Communication, detail how "CTV networks and programming services have built far-reaching operations and partnerships to maximize the harvesting of data in order to serve the interests of advertisers."[33] The absence of sufficient government regulations against data collection and the inability for consumers to adequately opt-out of this collection, along with personalized advertising on connected television through the use of generative artificial intelligence, lead the pair to resolutely declare that "advertising and data collection are now the driving force in the connected television industry, shaping all of its operations, influencing its program offerings, and spawning a new generation of channels."[34] Chester and Montgomery's report identifies powerful industries, such as the pharmaceutical and food and beverage industries (alongside various political groups), who have significant interests in this consumer data, and who use it to generate hyper-targeted advertisements.[35] The surveillance threat posed by these televisions is so significant that Chester and Montgomery describe the purchasing of a smart TV as "akin to bringing a digital Trojan Horse into one's home."[36] Unfortunately, these smart TVs, and streaming in general, have become the "dominant way that people get television in the U.S."[37] Because of the monopolistic dominance of streaming giants like Netflix and Disney, alongside massive TV manufacturers like Roku, Samsung, and LG, anyone

31 ▪ For these uses, see the entry for "publico" in Charlton T. Lewis and Charles Short, *A Latin Dictionary* (Clarendon Press, 1879).

32 ▪ For a historical overview on the rise of platforms, see Nick Srnicek, *Platform Capitalism* (Polity Press, 2016). See also Rebecca Giblin and Cory Doctorow's *Chokepoint Capitalism* (Beacon Press, 2023) for a wider critique of the "chokepoints" in capitalism which enable corporate monopolies and monopsonies like Google, Amazon, and Apple, to strangle competition through and beyond platforms. Lastly, see Sarah Lamdan, *Data Cartels: The Companies that Control and Monopolize Our Information* (Stanford University Press, 2022) for a larger critique of "data brokers."

33 ▪ Jeff Chester and Kathryn C. Montgomery, *How TV Watches Us: Commercial Surveillance in the Streaming Era* (Center for Digital Democracy, 2024), 6.

34 ▪ Chester and Montgomery, 6.

35 ▪ See Chester and Montgomery, 26-7 for a discussion of the pharmaceutical and food and beverage industries' interests in connected television data and advertising.

36 ▪ Chester and Montgomery, 24.

37 ▪ Chester and Montgomery, 5.

wanting to view the most popular contemporary cultural productions (or even the daily news) on connected television must also subject themselves to an intrusive data collection and advertising regime.

Publishing proper—that is, public distribution by way of digital ownership and physical media—allows the public the chance to build and preserve their own cultural records beyond corporate and government control, and in some instances, even beyond the control of large public-serving archives.[38] For example, VHS and DVD distributions historically allowed 'cult followings' to emerge around unorthodox films like *Dazed and Confused, Donnie Darko, and Fight Club*, causing these films to be critically reevaluated as culturally significant long after they originally flopped at the box office.[39] That same process of an independent and passionate public keeping subcultural expression alive cannot happen with media that the public cannot own. Without physical distribution or digital ownership for individuals and libraries, digital content such as ebooks and audiobooks, streaming-only music, film, television shows, online news, licensed software, and more can completely disappear when licenses expire and when publishers decide that it is no longer profitable (economically and ideologically) to keep that content online.[40] And although content taken offline *might* exist in a publisher's private archive, there is no guarantee that it will ever be available to the public again.[41]

38 ▪ See e.g. Mark O'Brien's contribution to this report, "On Filmstrips" for a discussion of non-institutional preservation efforts.

39 ▪ For a discussion of these films and others, see Emily Barker, "13 box office flops that became hugely successful on DVD," *NME*, October 21, 2015.

40 ▪ For a succinct explanation of how licensing issues cause streaming content to disappear from public access, see Sarah Whitten and Lillian Rizzo, "Streaming services are removing tons of movies and shows — it's not personal, it's strictly business," CNBC, May 29, 2023. Note that publishers do not go around publicizing how much of their content is no longer available. Third-party sites frequently document these removals, see e.g. Roger Palmer, "What's Leaving Hulu & Hulu On Disney+ In September 2024," What's on Disney Plus, August 19, 2024 and Kasey Moore, "Every Netflix Original Series and Movie Removed from Netflix," *What's on Netflix*, July 26, 2024. More quantitative research is needed to understand the extent of these losses.

41 ▪ Film, television, and new media scholar Mikhail Skoptsov argues in a Substack post titled "Why Removed HBO Max Originals Return in Other Formats" that media scholars ought to approach the issue of cultural 'loss' from the point of cultural reissuing. Surveying HBO Max Originals removed from the platform from July 2022 to October 2023. Skoptsov found that of the forty-eight Originals removed from the platform, thirty of them (62.5%) had been reissued either on a different streaming platform or in a different format. Nevertheless, eighteen of the shows remained unavailable, and there is no certainty that these Originals will stay on their different platforms for long.