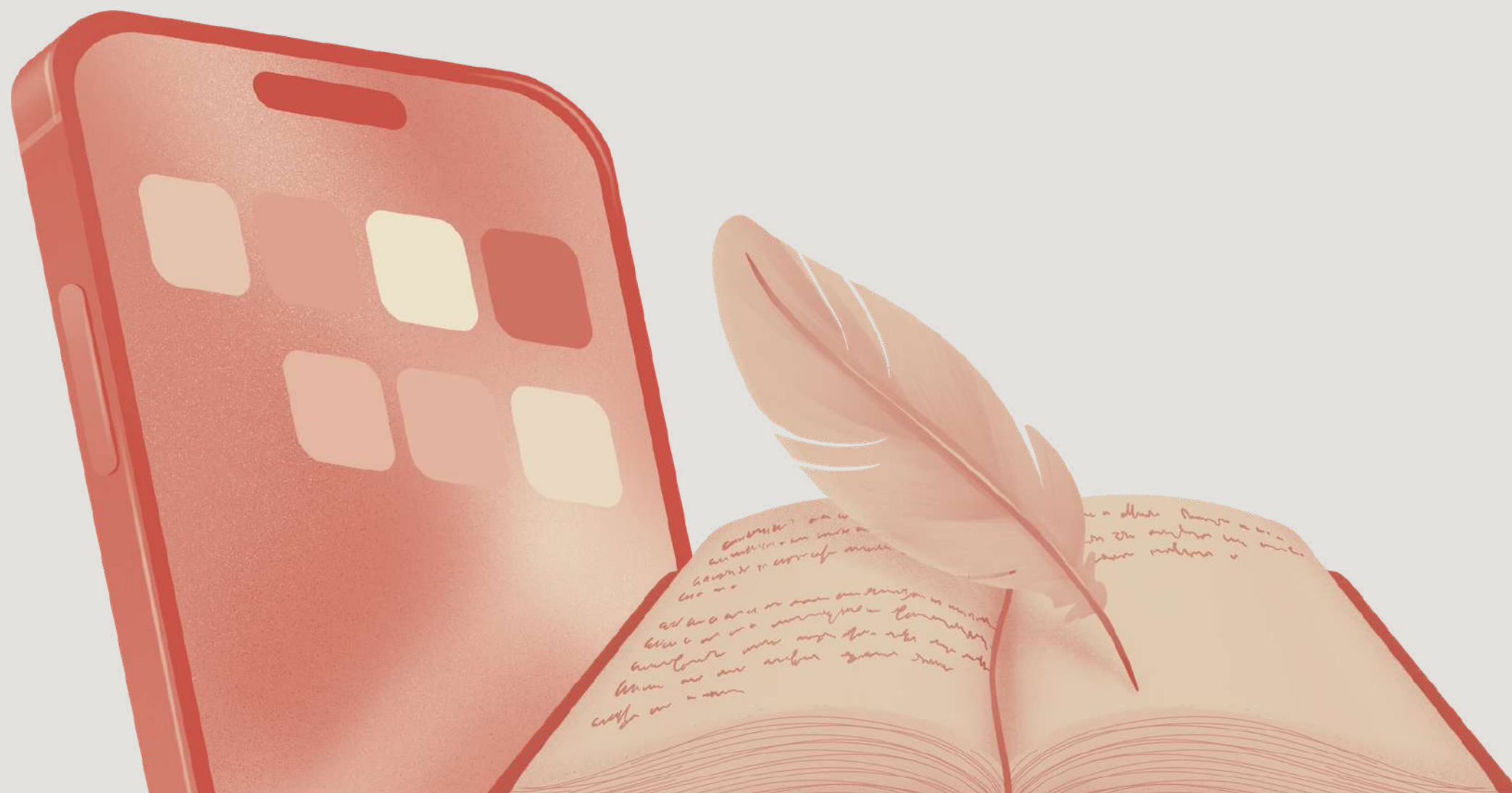


Инфраструктура забывания

*Технологические и онтологические
уязвимости веб-архива WayBack Machine*



ДЕНИСОВА ВАРВАРА
СЕМИНАР ИТМО X ВШЭ
18 ИЮНЯ 2026

Кейс



[DONATE](#)

INTERNET ARCHIVE
WayBackMachine

Explore more than 1 trillion [web pages](#) saved over time



**Tools**

[Wayback Machine Availability API](#)
[Chrome Extension](#)
[Firefox Add-on](#)
[Safari Extension](#)
[MS Edge Add-on](#)
[iOS app](#)
[Android app](#)

**Subscription Service**

Archive-It enables you to capture, manage and search collections of digital content without any technical expertise or hosting facilities.[Visit Archive-It to build and browse the collections.](#)

**Collection Search**

This service is based on indexes of specific data from [selected Collections](#).

**Save Page Now**

Capture a web page as it appears now for use as a trusted citation in the future.

[FAQ](#) | [Contact Us](#) | [Terms of Service \(Dec 31, 2014\)](#)

Введение



Совокупный объем мировых данных достиг 180 зеттабайт в 2025 году, ежегодный рост — 23% (Statista, 2025). Объем WBM составляет около 99 петабайт, что эквивалентно ~1 триллиону сохраненных веб-страниц (Internet Archive, 2025).



WBM был создан для обеспечения «постоянного доступа» к эфемерному веб-контенту (Archive-It, 2024) — как библиотека Интернета, заявляющая устойчивость и стабильность веб-памяти.



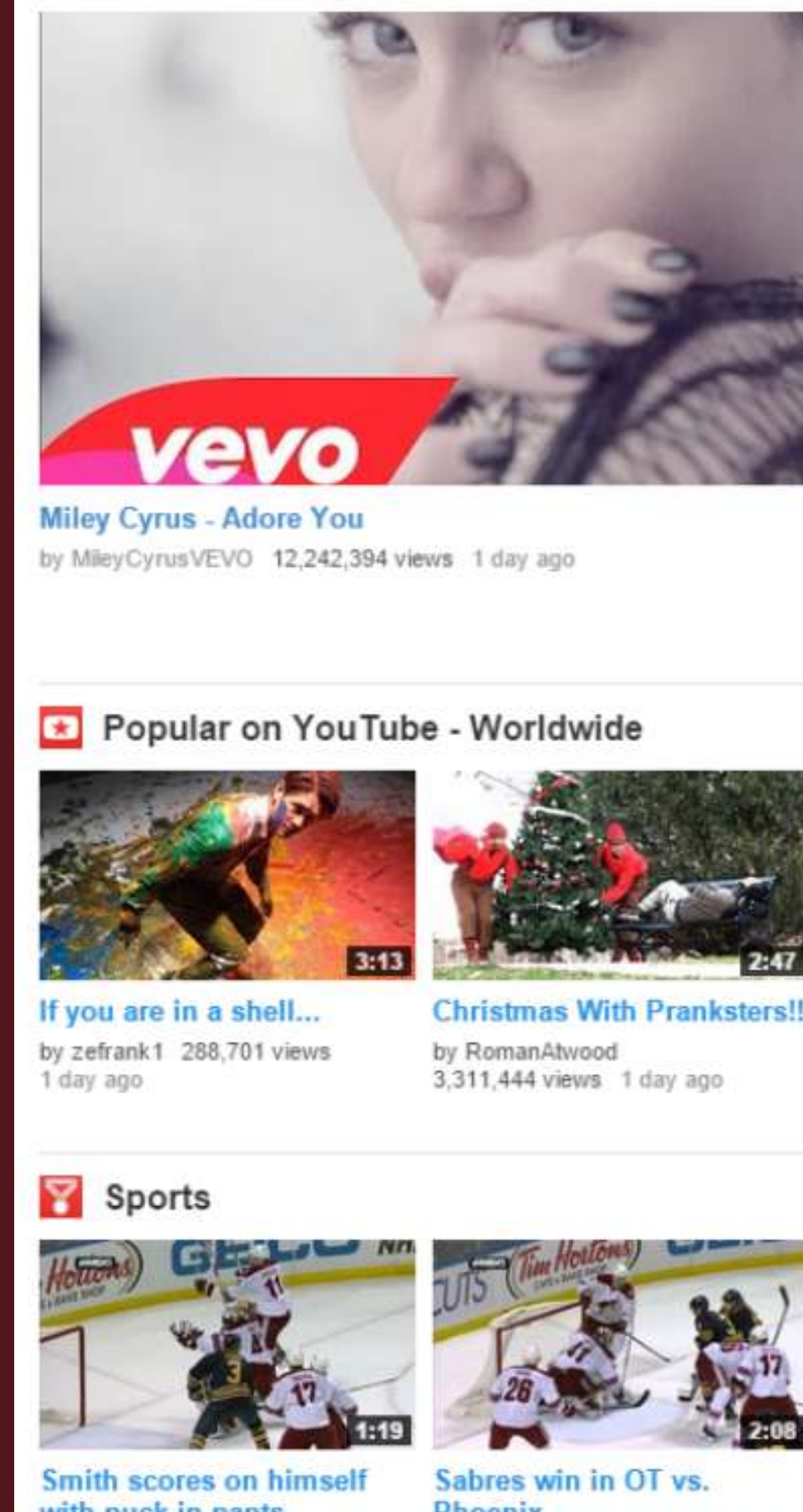
Однако архив неизбежно селективен, фрагментарен и зависит от инфраструктурных ограничений. «Постоянство» противоречит онтологическому качеству конечности (Derrida, 1996, с. 93).

Данное исследование анализирует забывание как процесс, формируемый цифровой инфраструктурой WBM, — рассматривая технологическую уязвимость как трансформирующее условие, порождающее новые архивные формы.

Прошлый ресерч

Betrayed Memory: The Death of Information in Web Archives (Denisova, 2025)

- Память WBM не является ни полностью сохранённой, ни полностью утраченной — она гипомнестична: частичная, постепенная утрата (Denisova, 2025; Kyong Chun, 2011, с. 195).
- Контент WBM представляет собой «монумент» в фукоянском смысле (Foucault, 1969, с. 8).
- Битые ссылки нарушают темпоральное понимание (Schlick, 1974, с. 189), тогда как кнопки Inspect/About индексируют контраст с «живым» потоком веба.
- WBM воплощает «влечение к смерти» Интернета (Freud, 1920): он работает против веба, а не сохраняет его (Denisova, 2025).
- Исследование смещает фокус с содержания на инфраструктурный контейнер памяти.



Разрыв

- В материальных архивах исчезновение было идентифицируемо (Vismann, 2008).
- В WBM уязвимости не идентифицируются: сбои серверов, повреждение баз данных, устаревание ПО, алгоритмическая фильтрация и медиаполитические решения формируют забывание без явных маркеров.
- И люди (пользователи), и нечеловеческие агенты (краулеры) распределяют процесс забывания, делая его инфраструктурно распределённым и предзаданным машинной онтологией.

Как зафиксировать процессы забывания, порождаемые алгоритмической работой Wayback Machine, и тем самым цифровыми архивами?



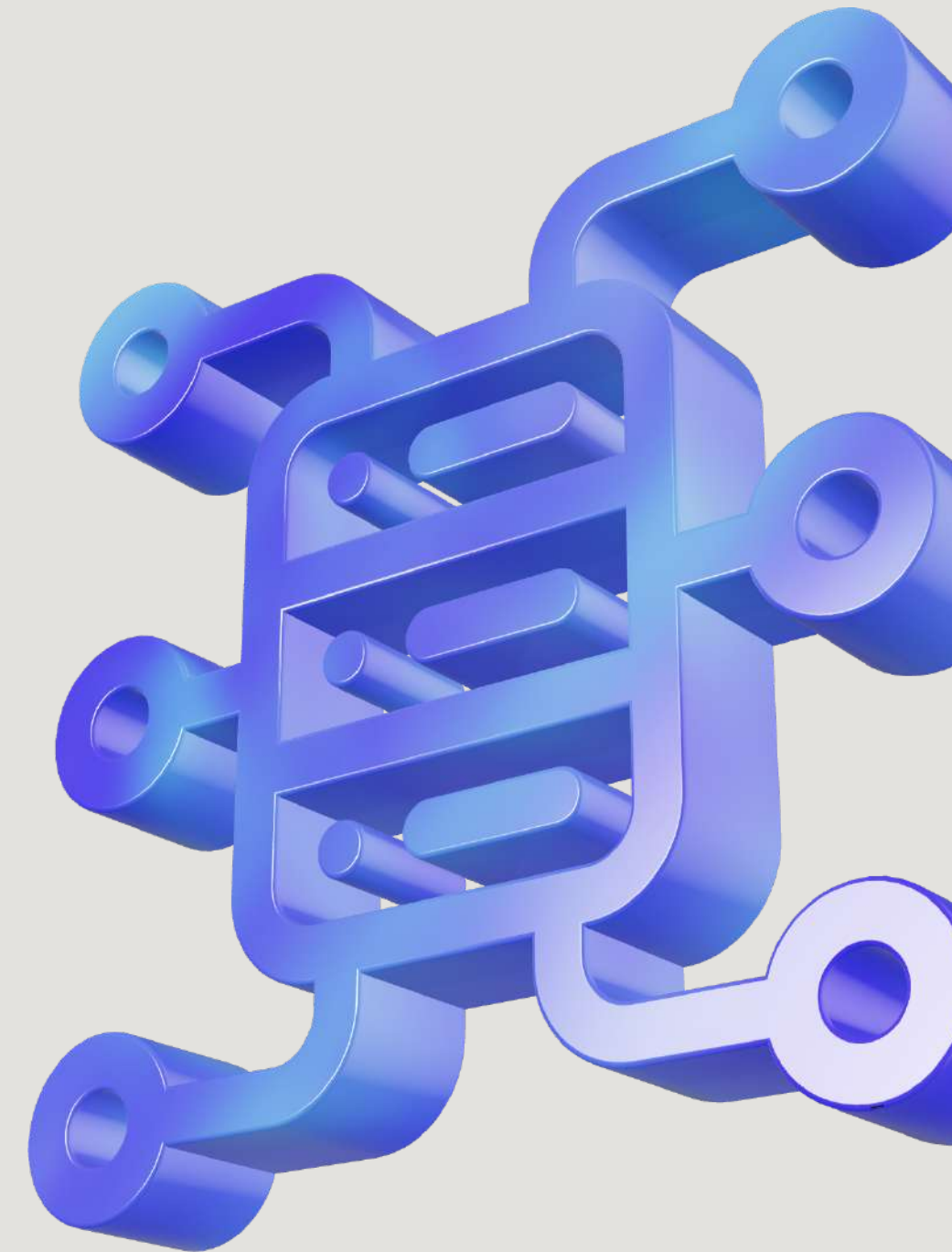
Гипотеза

Цели

Забывание в веб-архивах является конститутивной и заранее предзаданной операциональной характеристикой машинных инфраструктур, организующих цифровую память.

Реконфигурация памяти в веб-архивах определяется алгоритмическими логиками и инфраструктурными ограничениями.

- 1 В данном исследовании WBM концептуализируется как инфраструктурная конфигурация, в которой забывание со-производится взаимодействием человеческих и нечеловеческих акторов.
- 2 Исследование разрабатывает методологический каркас для отслеживания того, как технические операции заранее программируют забывание как детерминированное событие внутри веб-памяти.



Literature review

5

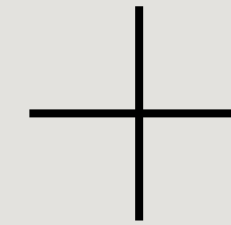
Философия техники

Анализ материальной инфраструктуры

Дискурс анализ

Технологические
Онтологические

уязвимости



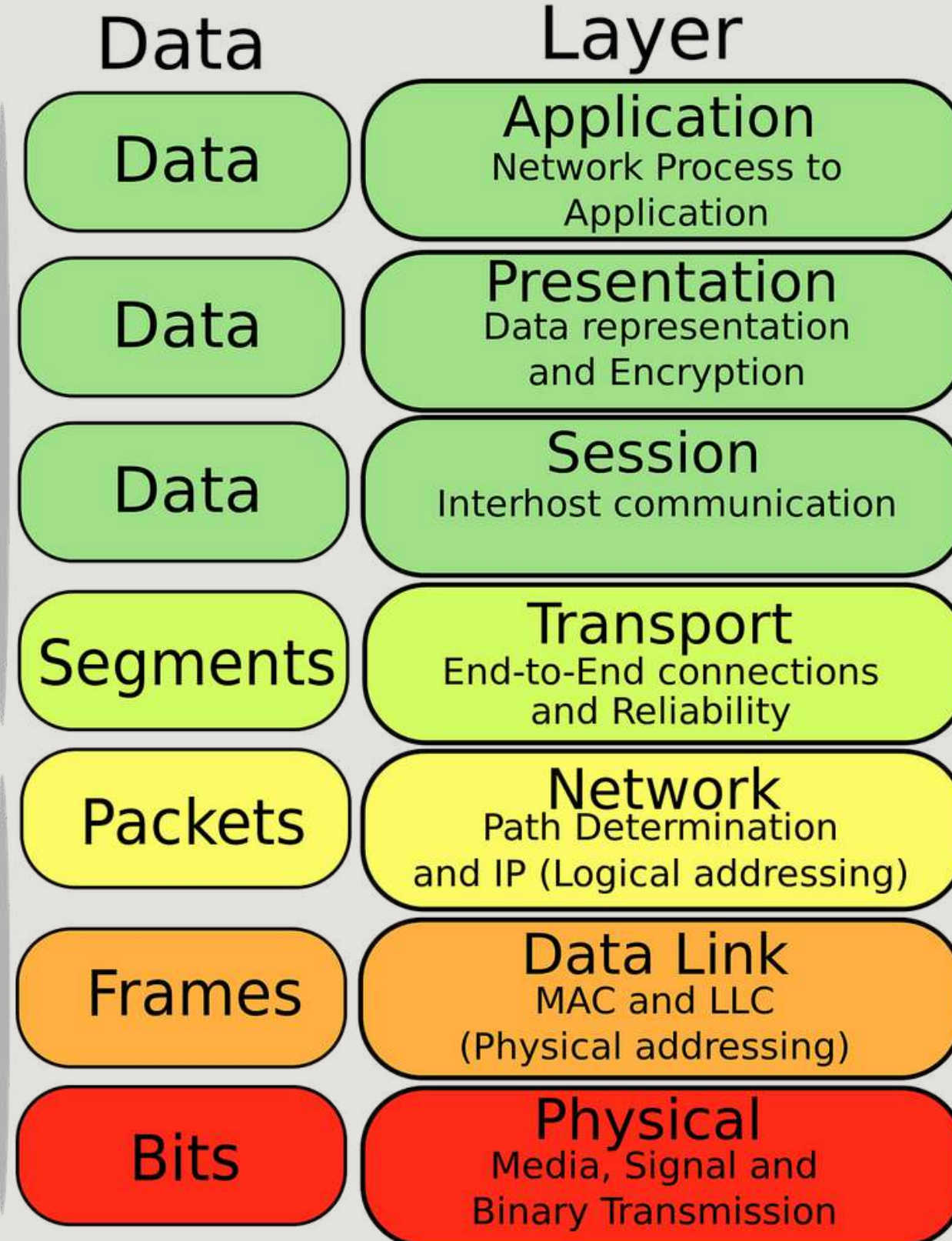
Цифровая инфраструктура

RQ

OSI Model

Host Layers

Media Layers



HTML, CSS, JS, шрифты,
медиа
форматы и кодировки

Забывание (формальная цифровая материальность):

- отсутствующие CSS/JS → «сломанный» интерфейс
- неархивированные медиа (из внешних CDN)
- несоответствие форматов, устаревшие стандарты

влияет

влияет

Уровень приложения
(OSI 7)

Веб-приложение WBM / SPN
логика захвата и воспроизведения

Забывание (аппликационный уровень):

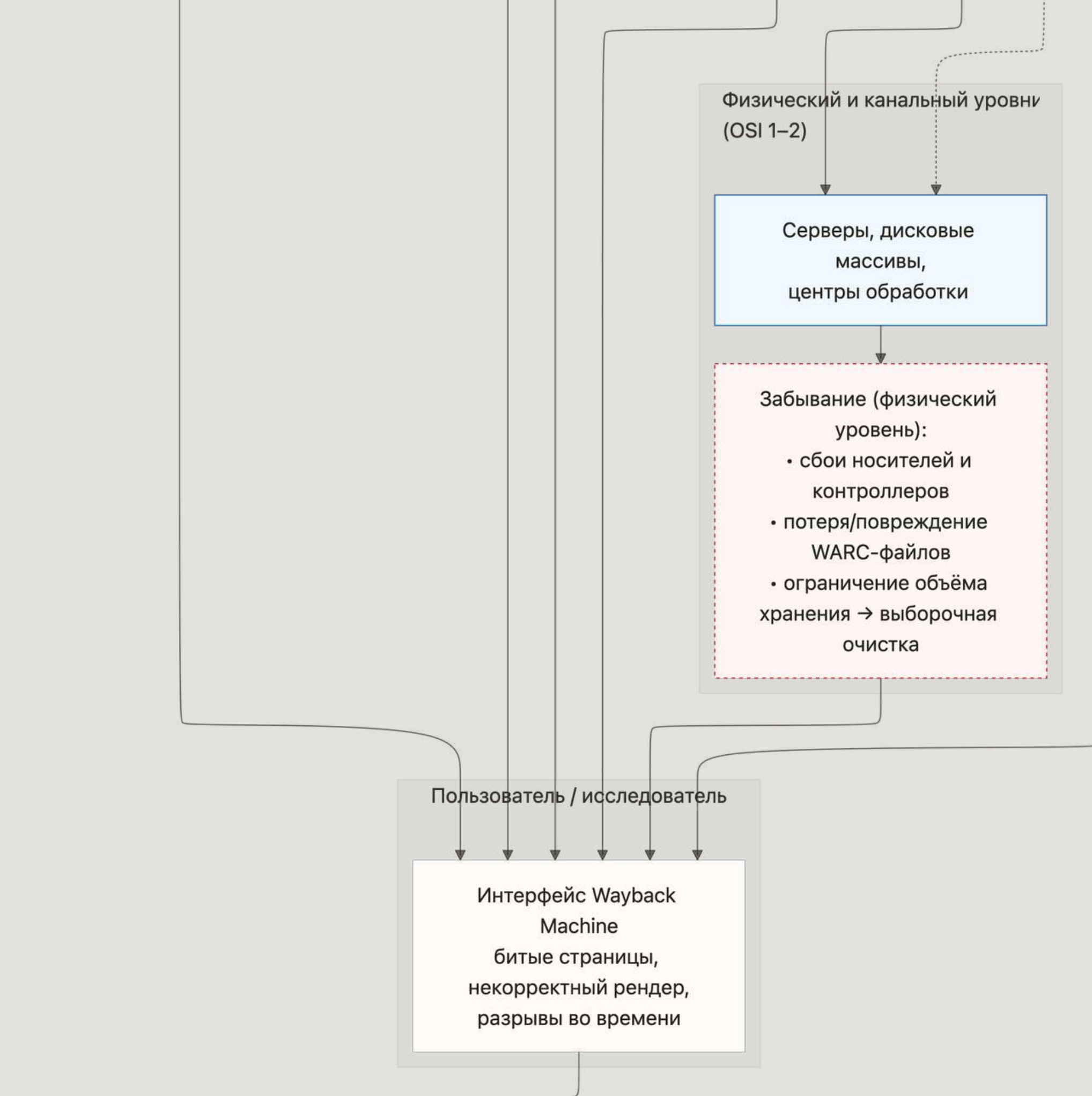
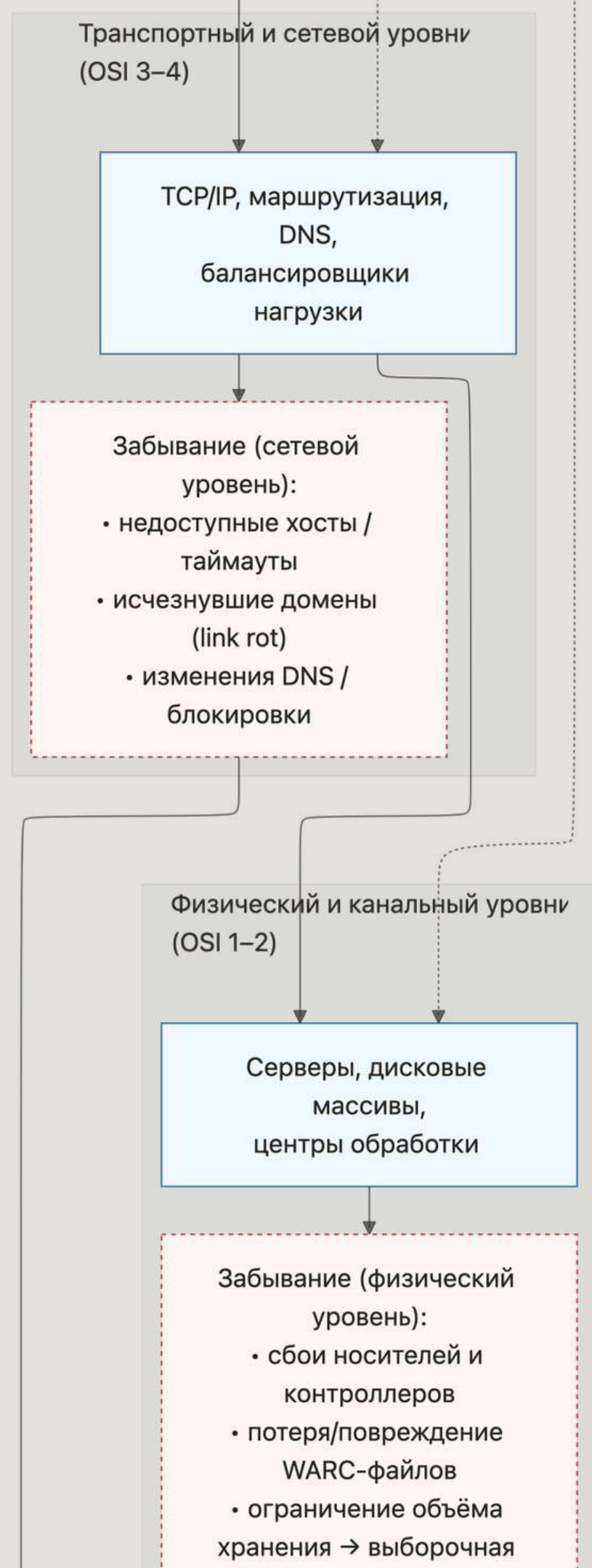
- отсутствующие снимки (calendar gaps)
- отсутствие доменов / исключения по политике
- манипулированные или удалённые снапшоты

Уровень сеанса и HTTP-логики
(OSI 5)

HTTP-сессии, куки, редиректы, заголовки, Content Negotiation

Забывание (сеансовый уровень):

- бесконечные/ошибочные редиректы
- запросы к «живому» вебу вместо архива
- потеря части ресурсов из-за особенностей Negotiation



The ~~Online~~ Offline Archive

 INTERNET
ARCHIVE

 WEB

 TEXTS

 VIDEO

 AUDIO

 SOFTWARE

 IMAGES

 SIGN UP | LOG IN

 UPLOAD



ABOUT

BLOG

EVENTS

PROJECTS

HELP

DONATE 

CONTACT

JOBS

VOLUNTEER

The Offline Internet Archive

The internet now seems like a utility, available everywhere from our homes and offices to trains and planes. But utility-level access is not yet a reality for more than half of the world's population who lack consistent, or indeed any access, to the Internet.

Why?

- Cost: Internet access is unaffordable to people with low or no income.
- Connectivity: In many developing countries and rural areas the infrastructure that enables internet access is unreliable, slow, or simply unavailable. Natural disasters, uprisings, and war compound the challenge.
- Censorship: Some countries limit internet access for political reasons. Several countries block the Internet Archive. In some countries, Facebook has become synonymous with the internet – but it is hardly a substitute for free and open World Wide Web.

The Internet Archive offers perhaps the world's largest online store of open content. The wisdom of the ages, just a few clicks away. As Wikipedia has become the world's encyclopedia, the Internet Archive has become its library. Central to our mission is establishing "Universal Access to All Knowledge". Access to our library of millions of books, journals, audio and video recordings and beyond is free to anyone — with one caveat — the need for a reliable internet connection.

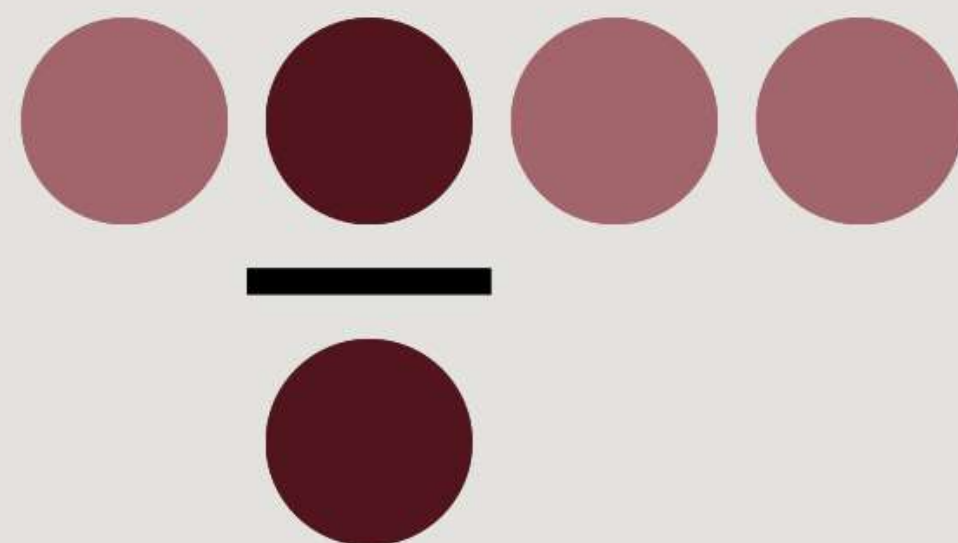
Lack of access to today's internet is a significant factor in poorer educational outcomes, inter-generational poverty and disempowerment as identified by the UN in their Sustainable Development Goal #9. The Offline Archive project works towards making online collections available — regardless of internet availability.

Part of the challenge is that those of us who live where the Internet works well, are adding graphics, video and other demands on bandwidth faster than access is being improved in many parts of the world.

дискурс теория

ЭРНЕСТО ЛАКЛО И ШАНТАЛЬ МУФФ

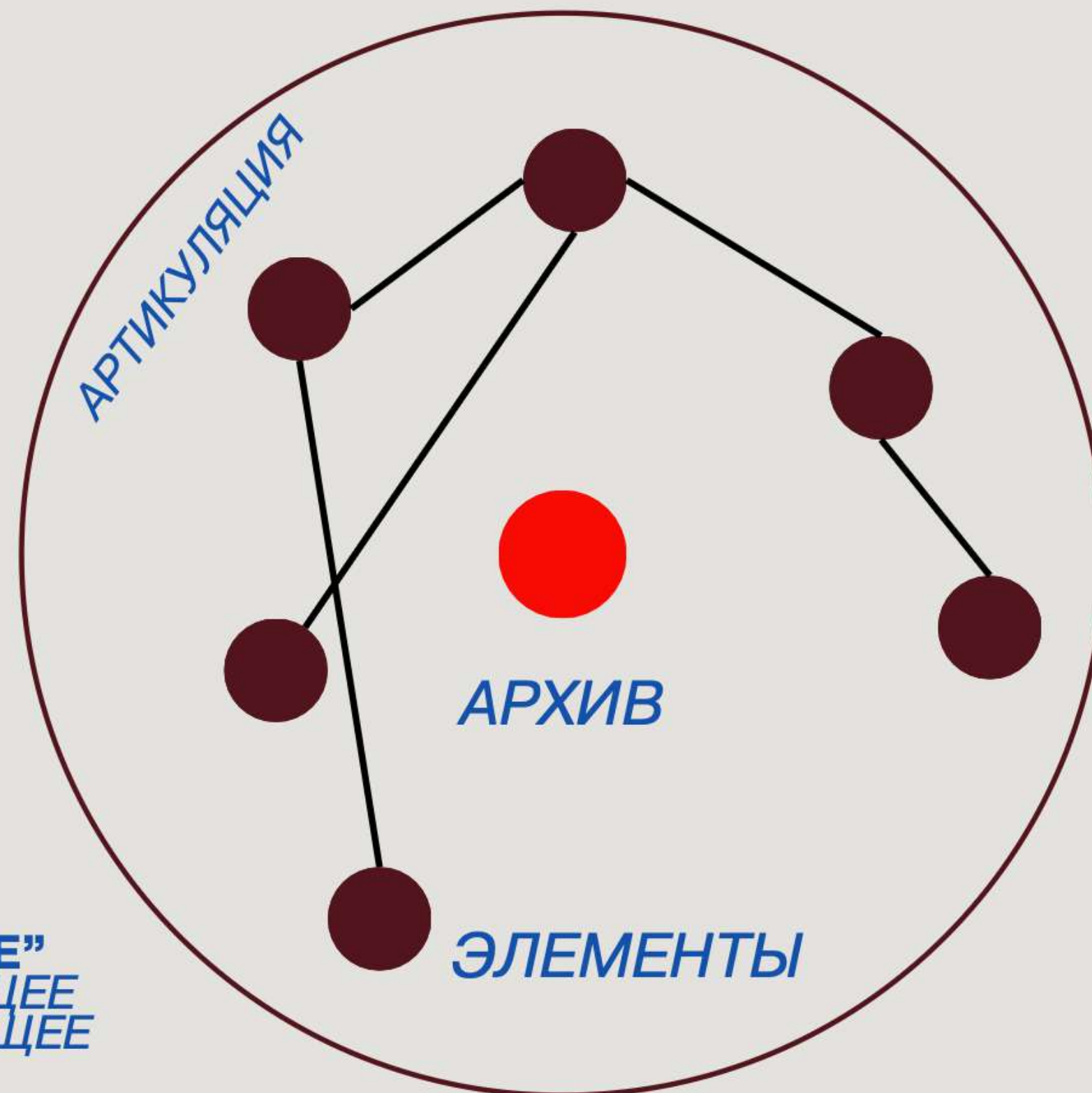
ДИСКУРС



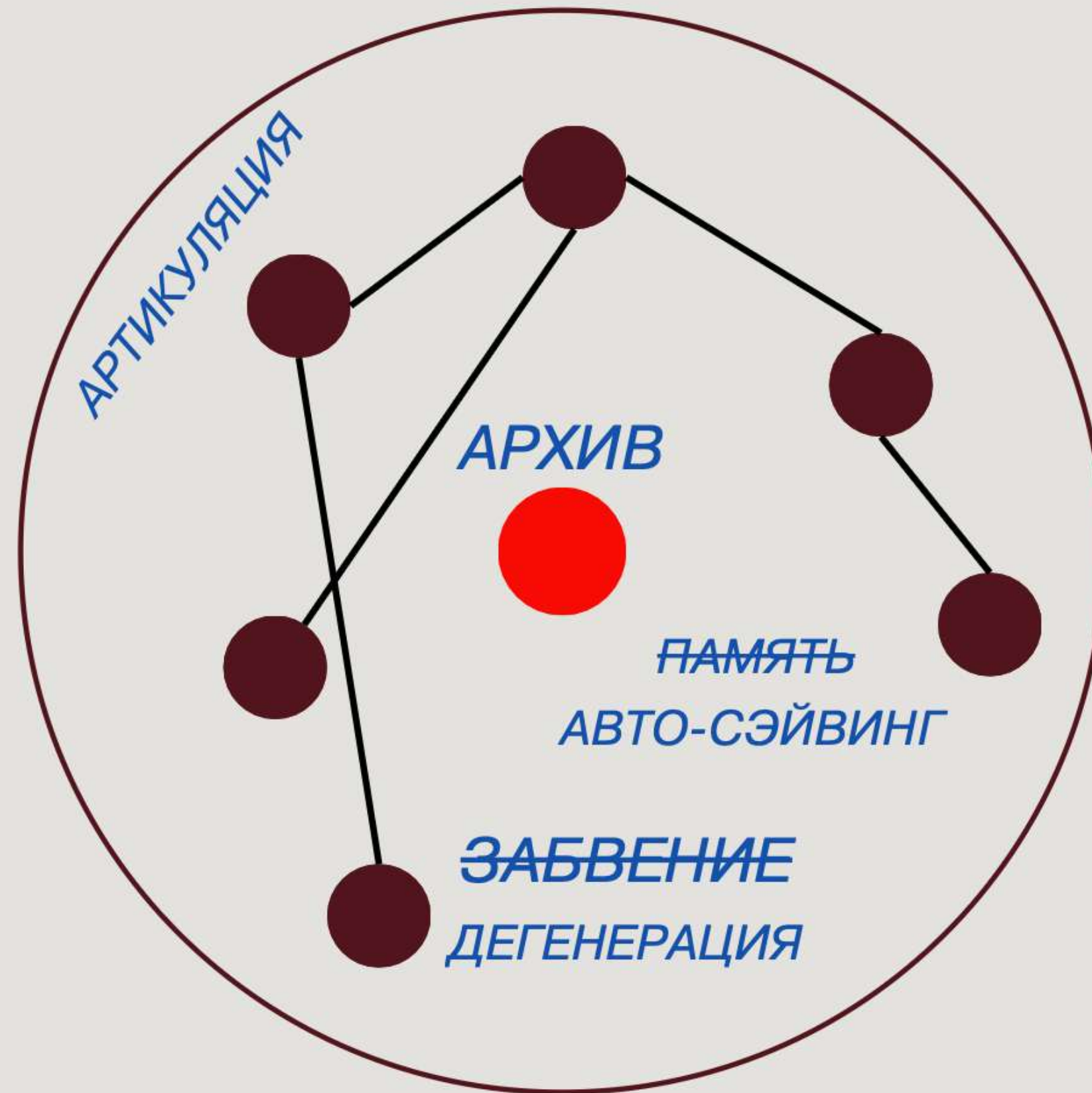
ОБРЕТАЕТ СМЫСЛ
В ОТНОШЕНИИ К
ИНОМУ МОМЕНТУ



“OFFLINE”
ПЛАВАЮЩЕЕ
ОЗНАЧАЮЩЕЕ



ДИСКУРС



“Offline archives” are articulated as solutions to digital vulnerability

 INTERNET
ARCHIVE

 WEB

 TEXTS

 VIDEO

 AUDIO

 SOFTWARE

 IMAGES

 SIGN UP | LOG IN

 UPLOAD

 Search

ABOUT

BLOG

EVENTS

PROJECTS

HELP

DONATE 

CONTACT

JOBS

VOLUNTEER

The Offline Internet Archive

The internet now seems like a utility, available everywhere from our homes and offices to trains and planes. But utility-level access is not yet a reality for more than half of the world’s population who lack consistent, or indeed any access, to the Internet.

Why?

- Cost: Internet access is unaffordable to people with low or no income.
- Connectivity: In many developing countries and rural areas the infrastructure that enables internet access is unreliable, slow, or simply unavailable. Natural disasters, uprisings, and war compound the challenge.
- Censorship: Some countries limit internet access for political reasons. Several countries block the Internet Archive. In some countries, Facebook has become synonymous with the internet – but it is hardly a substitute for free and open World Wide Web.

The Internet Archive offers perhaps the world’s largest online store of open content. The wisdom of the ages, just a few clicks away. As Wikipedia has become the world’s encyclopedia, the Internet Archive has become its library. Central to our mission is establishing “Universal Access to All Knowledge”. Access to our library of millions of books, journals, audio and video recordings and beyond is free to anyone — with one caveat — the need for a reliable internet connection.

Lack of access to today’s internet is a significant factor in poorer educational outcomes, inter-generational poverty and disempowerment as identified by the UN in their Sustainable Development Goal #9. The Offline Archive project works towards making online collections available — regardless of internet availability.

Part of the challenge is that those of us who live where the Internet works well, are adding graphics, video and other demands on bandwidth faster than access is being improved in many parts of the world.

An evolving ecosystem is emerging to enable access over poorer internet. Typically the approaches build around low cost, low power, devices that can be installed, in communities and schools for example, and deliver content either offline or through better usage of a narrow pipe to the net.

We have built an offline server that:

- Crawls Internet Archive collections to a local server.
- Serves that content locally,
- Caches content while browsing.
- Moves content between servers by sneakernet — on disks, USB sticks, and SD cards.
- Delivers (mostly) the Internet Archive UI offline in javascript in the browser,
- Is open source
- And is being made available in other languages.

Список литературы

- Archive-It. (2024). About Archive-It. <https://archive-it.org/learn-more/>
- Cambridge Dictionary. (2025). Anachronism. Cambridge University Press & Assessment. <https://dictionary.cambridge.org/dictionary/english/anachronism>
- Cartwright, L. (1995). Experiments of destruction: Cinematic inscriptions of physiology. In L. Cartwright (Ed.), *Screening the body: Tracing medicine's visual culture* (pp. 17–46). University of Minnesota Press.
- Denisova, V. (2025). Betrayed memory: The death of information in web archives (Bachelor's thesis, Tyumen State University, School of Advanced Studies).
- Derrida, J. (1996). *Archive fever: A Freudian impression* (E. Prenowitz, Trans.). University of Chicago Press.
- Foucault, M. (1969). *The archaeology of knowledge and the discourse on language*. Pantheon Books.
- Freud, S. (1920). *Beyond the pleasure principle*. The International Psycho-Analytical Library.
- Fuchs, C. (2015). *Culture and economy in the age of social media*. Routledge.
- Garg, K., Alam, S., Ayala, D., Weigle, M. C., & Nelson, M. L. (2025, December). Resilience, volume, and temporal trends across 25 years of the Wayback Machine. In *2025 ACM/IEEE Joint Conference on Digital Libraries (JCDL)* (pp. 59–68). IEEE. <https://doi.org/10.1109/JCDL67857.2025.00017>
- Hagen, W. (2006). The style of sources. In W. H. Kyong Chun & T. Keenan (Eds.), *New media, old media: A history and theory reader* (pp. 157–175). Psychology Press.
- Hayles, N. K. (1999). *How we became posthuman: Virtual bodies in cybernetics, literature, and informatics*. University of Chicago Press.
- Internet Archive. (n.d.). Donate. <https://archive.org/donate>
- Jørgensen, M. W., & Phillips, L. (2002). *Discourse analysis as theory and method*. Sage.
- Kumar, B. T. S., & Prithviraj, K. R. (2015). Bringing life to dead: Role of Wayback Machine in retrieving vanished URLs. *Journal of Information Science*, 41(1), 71–81. <https://doi.org/10.1177/0165551514552752>
- Kumar, S., Dalal, S., & Dixit, V. (2014). The OSI model: Overview on the seven layers of computer networks. *International Journal of Computer Science and Information Technology Research*, 2(3), 461–466.
- Lacan, J. (2006). *Écrits* (B. Fink, Trans.). W. W. Norton.
- Lippit, A. M. (2005). *Atomic light (shadow optics)*. University of Minnesota Press.
- McPherson, T. (2006). Reload: Liveness, mobility, and the web. In W. H. Kyong Chun & T. Keenan (Eds.), *New media, old media: A history and theory reader* (pp. 199–208). Psychology Press.
- Momro, J. (2013). Derrida in the archive: Genetic spectres. *Czas Kultury*, 2, 84–97. <https://www.ceeol.com/search/article-detail?id=83245>
- Ogden, J., Summers, E., & Walker, S. (2024). Know(ing) infrastructure: The Wayback Machine as object and instrument of digital research. *Convergence*, 30(1), 167–189.
- Schou, J. (2016). Ernesto Laclau and critical media studies: Marxism, capitalism, and critique. *tripleC*, 14(1), 292–311.
- Schlick, M. (1974). The temporality of the real. In *General theory of knowledge (LEP Library of Exact Philosophy, Vol. 11, p. 189)*. Springer.
- Shep, S. J. (2015). Digital materiality. In S. Schreibman, R. Siemens, & J. Unsworth (Eds.), *A new companion to digital humanities* (Chap. 22). <https://doi.org/10.1002/9781118680605.ch22>
- Statista. (2025). Worldwide data created [Statistic]. Statista. <https://www.statista.com/statistics/871513/worldwide-data-created/>
- Vismann, C. (2008). *Files: Law and media technology* (G. Winthrop-Young, Trans.). Stanford University Press.
- Wendy Chun, H. K. (2011). The enduring ephemeral, or the future is a memory. In E. Huhtamo & J. Parikka (Eds.), *Media archaeology: Approaches, applications, and implications* (pp. 184–204). University of California Press. <https://doi.org/10.1525/9780520948518-011>