la rivista di **engramma**
marzo **2025**

222

Copyleft & internauti pirati

edizioni**engramma**

La Rivista di Engramma
222

La Rivista di
Engramma
222
marzo 2025

Copyleft & internauti pirati

a cura di
Filippo Perfetti e Giulia Zanon

edizioni**engramma**

direttore
monica centanni

redazione
damiano acciarino, sara agnoletto, mattia angeletti,
maddalena bassani, asia benedetti, maria bergamo,
elisa bizzotto, emily verla bovino,
giacomo calandra di roccolino, olivia sara carli,
concetta cataldo, giacomo confortin,
giorgiomaria cornelio, vincenzo damiani,
mario de angelis, silvia de laude,
francesca romana dell'aglio, simona dolari,
emma filipponi, christian garavello, anna ghiraldini,
ilaria grippa, roberto indovina, delphine lauritzen,
annalisa lavoro, laura leuzzi, michela maguolo,
ada naval, viola sofia neri, alessandra pedersoli,
marina pellanda, filippo perfetti, chiara pianca,
margherita piccichè, daniele pisani, stefania rimini,
lucamatteo rossi, daniela sacco, cesare sartori,
antonella sbrilli, massimo stella, ianick takaes,
elizabeth enrica thomson, christian toson,
chiara velicogna, giulia zanon

comitato scientifico
barbara baert, barbara biscotti, andrea capra,
giovanni careri, marialuisa catoni, victoria cirlot,
fernanda de maio, alessandro grilli, raoul kirchmayr,
luca lanini, vincenzo latina, orazio licandro,
fabrizio lollini, natalia mazour, alessandro metlica,
guido morpurgo, andrea pinotti, giuseppina scavuzzo,
elisabetta terragni, piermario vescovo, marina vicelja

comitato di garanzia
jaynie anderson, anna beltrametti, lorenzo braccesi,
maria grazia ciani, georges didi-huberman,
alberto ferlenga, nadia fusini, maurizio harari,
arturo mazzarella, elisabetta pallottino,
salvatore settis, oliver taplin

La Rivista di Engramma
a peer-reviewed journal
222 marzo 2025
www.engramma.it

sede legale
Engramma
Castello 6634 | 30122 Venezia
edizioni@engramma.it

redazione
Centro studi classicA Iuav
San Polo 2468 | 30125 Venezia
+39 041 257 14 61

©2025

edizioni**engramma**

ISBN carta 979-12-55650-80-5
ISBN digitale 979-12-55650-81-2
ISSN 2974-5535
finito di stampare maggio 2025

Si dichiara che i contenuti del presente volume sono la versione a stampa totalmente corrispondente alla versione online della Rivista, disponibile in open access all'indirizzo: https://www.engramma.it/222 e ciò a valere ad ogni effetto di legge.
L'editore dichiara di avere posto in essere le dovute attività di ricerca delle titolarità dei diritti sui contenuti qui pubblicati e di aver impegnato ogni ragionevole sforzo per tale finalità, come richiesto dalla prassi e dalle normative di settore.

Sommario

7 *Copyleft & internauti pirati*
Filippo Perfetti e Giulia Zanon

19 *"©Tutti i diritti riservati". Proprio tutti?*
Peppe Nanni

29 *Engramma Open Access. Aperta origine*
Redazione di Engramma

Invenzione

45 *Riprodurre il patrimonio culturale*
Mirco Modolo

69 *Per un'iconografia della noia*
Lorenzo Gigante

105 *A Tale of Two Misfits*
Maurizia Paolucci

117 *Una lettura di: Yan Thomas, Il valore delle cose, Quodlibet [2002] 2022*
Bernardo Prieto

Chiusura

131 *No logo?*
Giorgiomaria Cornelio

163 *Il potere segreto. Wikileaks e la digitalizzazione dell'informazione*
a cura di Alessandro Visca

173 *Museums and the Enclosure of the Public Domain in the Digital Age*
Douglas McCarthy

181 *Piracy Shield, diritto d'autore e monopoli intellettuali: il caso italiano*
Barbara Pasa

199 *Note su alcune aperture giurisprudenziali al principio del "libero utilizzo" e sugli effetti nelle culture del riuso*
Alessia Brandoni

Apertura

211 *Copyright and Copyleft nell'era dell'intelligenza artificiale*
Alessandro Del Ninno

225 *Contro il copyright*
Francesco D'Isa

243 *Anna, the Universal Library*
Elizaveta Kozina, Christian Toson

261 *Per una liberazione dalla retorica dell'archivio*
Alessandro Gagliardo con Filippo Perfetti e Giulia Zanon

Dalla cambusa

291 *"Corrispondenza con un editore"*
Internazionale Situazionista, con una Presentazione di Filippo Perfetti

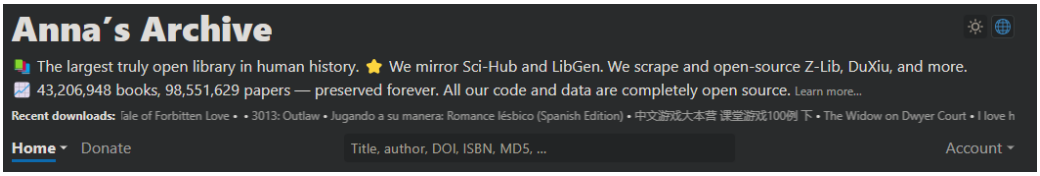
299 *Guerilla Open Access Manifesto*
Aaron Swartz, con una Presentazione di Giulia Zanon

305 *Elogio de la piratería*
Mauricio Souza Crespo, con una Presentazione di Filippo Perfetti

309 *10 tesi sull'archivio*
gruppo pad.ma, edizione italiana a cura della redazione di Inannunciato

Anna, the Universal Library

Elizaveta Kozina, Christian Toson



Who is Anna?
You are Anna!

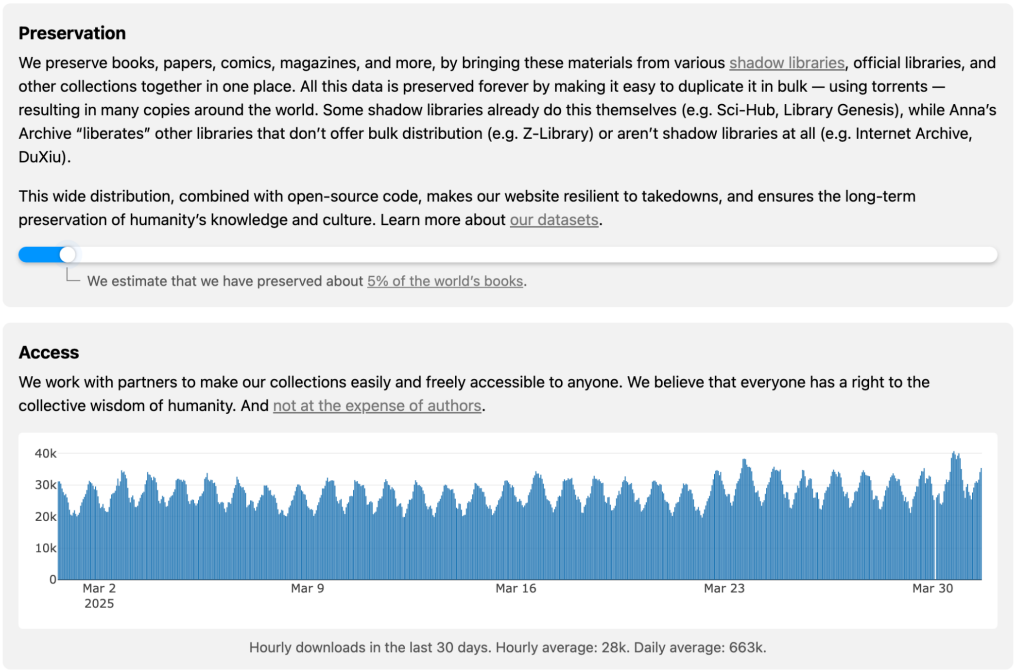
Archivists by name, librarians by self-definition, volunteers by activity, non-conformists by statements, pirates by necessity – Anna's team presents their project as "the largest truly open library in human history". Anna's library (or archive?), at the time of writing counts over 650,000 books downloaded every day. Taking a rough circulation estimate of the largest public library in the world, the New York Public Library, which amounts to around 30,000 copies per day, we notice that it accounts for only 10% of Anna's distribution. Currently Anna's Archive has collected more than 40 million books and around 100 million scientific papers. This should be enough significant data to deserve an investigation into her role in the global panorama of book circulation and copyright. It may be easy to dismiss Anna as petty piracy, a "little trick" that easy-going professors whisper as an aside to their students, but it is impossible to consider Anna as merely a deviation in our lives governed by rules. Anna undoubtedly deserves better treatment and recognition. Anna's Archive is an attempt to become a new kind of library, one that participates in the dialogue about knowledge and that seeks to change the rules of its circulation.

Anna's library: Utopian project or ordinary piracy?

We can approach the Anna phenomenon from two angles: through the description of her tasks and from the history of the so-called "shadow libraries". As for the first, Anna articulates the essence of her activity very clearly, since the team states that it is a non-profit project with two goals:

1) Preservation – backing up all knowledge and culture of humanity;
2) Access – making this knowledge and culture available to anyone in the world (https://annas-archive.org/faq).

As we can see, Anna's ultimate goal is not merely to found a public library, but to create a digital copy of existing books online (ideally, all books) in order to prevent them from disappearing.



A tracker keeps users updated in real time on how many books have already been copied. This material comes mostly from online repositories, both official and unofficial, which Anna calls "shadow libraries". In this sense, Anna's Archive is, in many ways, the latest instance in a long history of online libraries.

The term 'shadow library' has a double and contradictory definition. The first refers to large databases of written texts that are stored in non-open access platforms, such as, for example, the JSTOR library. 'Shadow', in this sense, means that the user can be aware of its existence, because texts can be found, for example, with a search tool, but, as in a shadow, their content is concealed. The second and most common definition of a shadow library is that of a large digital archive that collects and distributes written material (academic papers, books), and sometimes other media, violating the copyright laws of some countries. In this case, the term 'shadow' means 'illegal'; the term 'Black Open Access' (Black OA) is used with the same meaning. Following the first definition, perhaps the largest shadow library in the Internet is Google Books, which originated from a 2004 massive book digitization project involving five prominent academic libraries in the United States. In the first four years, it was intended to be a universal digital library, but by 2008, copyright issues blocked access, which is now subject to subscription, even for the same libraries that provided the books.

Another important project to mention in this context is Internet Archive, the forerunner of many modern shadow libraries, although Anna directly indicates that she does not associate it with

this term. Initially founded in 1996 to provide "Universal Access to All Knowledge", Internet Archive started from the assumption that a wide array of pages were written and produced through the web which nobody cared to preserve. This led to the creation of the Wayback Machine, a powerful tool still available online that records 'snapshots' of present and past web pages. Slowly, the archive began collecting books and other multimedia content, evolving into a new kind of library: most of the material was not available for download, but users could read it online. Thanks to a special borrowing system, the Internet Archive offered access to many books (according to official sources, at the beginning of 2022 it had more than 35 million books), while theoretically respecting copyright laws. Internet Archive aspires to be a legitimate online library. Nevertheless, in 2020, it was sued by four giants of the English-language publishing industry (Hachette, Penguin, HarperCollins, John Wiley), which greatly reduced the accessibility of its content. This case highlights how controversial the concept of lending is within the regime of current copyright law. While in a physical library a user is allowed to borrow any book, in a digital library this is forbidden if electronic copies are on sale. This outcome suggests that what matters is not a clear definition of accessibility and use of content, but whether it affects publishing companies' revenues. The Internet Archive, which has the ambition to be an innovative cultural institution, is struggling to find a recognised status within this heavily biased framework, constantly kept in check by shifting copyright regulations.

Parallel to these huge enterprises, which originate from Silicon Valley culture, there exists an equally relevant history of post-Soviet online libraries. In one of the best studies published on the topic of shadow libraries, *Shadow Libraries: Access to Knowledge in Global Higher Education,* Balázs Bodó reconstructs the history of the creation of such shared archives within Soviet culture and counterculture. Bodó suggests that Russian pirate libraries emerged from these enmeshed contexts:

> [...] Communist ideologies of the reading nation and mass education; the censorship of texts; the abused library system; economic hardships and dysfunctional markets; and, most importantly, the informal practices that ensured the survival of scholarship and literary traditions under hostile political and economic conditions (Bodó 2018, 33).

From the very beginning, Soviet copyright was very different from its Western counterpart: the issue of securing revenue for private individuals was a nonexistent problem, and all publishers, film studios, etc., were state-owned. That meant that most books and movies where openly accessible or available at a very low price, as they still are today. Secondly, the great laboratory of Soviet counterculture developed a series of practices to distribute censored material, ranging from self-publication (*samizdat*), to the sharing of libraries and DIY methods of copying and distribution. The arrival of the Internet saw a continuation of those practices. The first shared online archives were set up by Russian-speaking scientists and users, who collectively digitized books by hand, downloaded the archives of the research institutions where they worked, and scoured the internet to format and organize libraries that would reach tens of thousands of titles, such as: Lib.ru (1994) by Maxim Moshkov, Kolkhoz (2002), Monoscop, Gigapedia (2000s). All those were to some degree related to each other. Then came

the widespread Library.ru (2010), that was sued in 2012 and shut down, and Library Genesis (2007), which inherited the dataset of Library.ru. In short, it can be said that Soviet counter-culture, along with the post-Soviet scarcity of well-stocked official distributors, influenced the development of what would later become pirate libraries, exploring and advancing the collection, organisation, and wide dissemination of knowledge by taking advantage of the loose copyright enforcement in the early years of the World Wide Web.

Anna is in many ways a continuation of a long line of shadow libraries, and she openly declares that she was inspired by Library Genesis. However, she significantly raises the ambition of her goals. LibGen does not aim to contain everything; its boundaries are shaped thorugh dialogue with the community, measured by the act of actively digitizing and sharing books while maintaining a low profile (Bodó 2018, 36). Anna's Archive, by contrast, seeks to create a universal library that includes and preserves all the books of the world, and does so bravely. In other words, Anna proposes a project with no equal in terms of scale and ambition. In this context, Anna is perhaps the one of the few online libraries, regardless of legal status, pursuing such goals on a completely non-profit basis. The team's work is mostly voluntary, and funding comes from user donations and paid memberships that increase download speed. Anna's budget is probably much lower than that of any major library in the world. However, she has learned from LibGen that preserving intellectual culture, in this case, books, in today's aggressive legal and digital environment requires creating the largest possible number of copies. This way, if one library is lost, copies will still exist elsewhere. These copies of the archive are called 'mirrors'. Users can not only download individual books but also create a complete replica of Anna, by copying its open source code. As of today, humanity's largest library requires one petabyte (one million of Gigabytes) of memory, roughly the same amount of data generated in a month of shooting of a Netflix series.

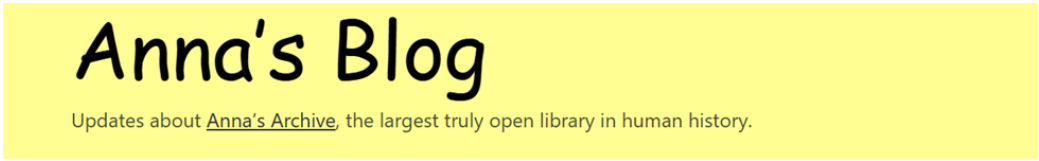
Today, alongside Anna, there is another system that, in some way, can be traced back to the the post-Soviet data sharing experience, a feature built into the messaging service of

Russia's most popular social network: VKontakte. Essentially a Russian version of Facebook, VKontakte has gained a reputation as one of the most reliable platforms for piracy, using a system for circulating private files. In other words, even a book or article sent in private messages can be found by a third-party user through the file search function. Thus, the process of uploading data is practically unregulated and does not rely on the efforts of a small team of volunteers, as in the case of Anna's Archive, but becomes a continuous, collective activity carried out by the network's users themselves. Moreover, all attempts to control it are nearly impossible, as this form of data circulation is spontaneous. At times, the collective forces of millions of users overcomes the restrictions imposed from above.

There inevitably arises the question of copyright. In recent decades, copyright laws have strictly determined the conditions of use and access to information and creative works, with little or no regard for the goals pursued by a particular project or its funding. The meaning of the term 'shadow library' in this context becomes more than just a label for the phenomenon of piracy; it signifies a form of resistance to outdated copyright standards and traditional concepts of what a library should be. For Anna 'shadow' is, in a sense, analogous to the late Soviet 'underground.' Anna is the 'underground' of copyright, a domain whose norms seek to make her cease to exist. This becomes clear when we read the very first post of the emerging Anna, when it wasn't yet Anna, but the "Pirate Library Mirror," in June 2022:

> **Introducing the Pirate Library Mirror (EDIT: moved to Anna's Archive): Preserving 7TB of books (that are not in Libgen)**
> This project aims to contribute to the preservation and liberation of human knowledge. We make our small and humble contribution, in the footsteps of the greats before us.
> The focus of this project is illustrated by its name:
> Pirate – We deliberately violate the copyright law in most countries. This allows us to do something that legal entities cannot do: making sure books are mirrored far and wide;
> Library – Like most libraries, we focus primarily on written materials like books. We might expand into other types of media in the future;
> Mirror – We are strictly a mirror of existing libraries. We focus on preservation, not on making books easily searchable and downloadable (access) or fostering a big community of people who contribute new books (sourcing)(https://annas-archive.org/blog/blog-introducing.html).

The Pirate Library Mirror was a highly technical title that synthesized the foundations upon which Anna was built. From the beginning, Anna was primarily conceived as a tool in open disobedience of copyright laws, which were seen as dangerous to the higher value of preserving knowledgem, hence the appellative of "pirate by necessity". "Library" refers simply to the fact that the focus of preservation is on written materials and books. This rebelliousness is carried out through mirroring: the creation of redundant copies that make it difficult to track hosting services and allow for resilience in the event of server shutdowns, hence 'mirror'. This method, however, meant that the system had to remain lightweight, both in terms of services (i.e., no tools for interacting with the data itself, such as search functions) and in terms of user interaction (i.e., no community network), features that are, for example, present in the Internet



Hi, I'm Anna. I created Anna's Archive, the world's largest shadow library. This is my personal blog, in which I and my teammates write about piracy, digital preservation, and more.

Connect with me on Reddit.

Note that this website is just a blog. We only host our own words here. No torrents or other copyrighted files are hosted or linked here.

Archive and are now almost considered standard on the modern Internet. Nevertheless, it is precisely Anna's simplicity and rawness that make it unique and worthy of more attention.

Anna's rhetoric: Simple and radical

Anna's Blog features minimal graphics, a yellow header, and a large Comic Sans title, reminiscent of the early Internet blogs from the 1990s. There are only a few entries from June 2022 to the present, which serve as the best sources for tracing the development of the project. Everything, from the visual design to the writing style, conveys that Anna aims to be simple, focused on a few primary tasks, and straightforward. Anna replaced the less appealing and anonymous Pirate Library Mirror. This change in name tells us something about the importance of identity:

> One decision to make for each project is whether to publish it using the same identity as before, or not. If you keep using the same name, then mistakes in operational security from earlier projects could come back to bite you. But publishing under different names means that you don't build a longer lasting reputation. We chose to have strong operational security from the start so we can keep using the same identity, but we won't hesitate to publish under a different name if we mess up or if the circumstances call for it.
>
> In some way, the name is the only thing that keeps the library unified, a holdfast. Books and files can be mirrored everywhere, but if we want to make a statement, at the end, it is necessary to fix a point, create an identity that can engage in a debate.

The necessity of a recognizable name, despite concerns about visibility and security, reveals that Anna aims to make a mark in the history of knowledge circulation. As they say, it is a "holdfast," a "fixed point," an "identity" that can initiate debate. The name Anna is biblical; also, is the name of the queen who, in 1710, introduced copyright (it is also a palindrome) but it could also represent the average name of an everyday reader, much like Caroline did for the journal Queen, which later inspired the name of the famous offshore pirate station, Radio Caroline (Pedersoli, Toson 2020). The name itself raises the question: "Who is Anna?", cleverly

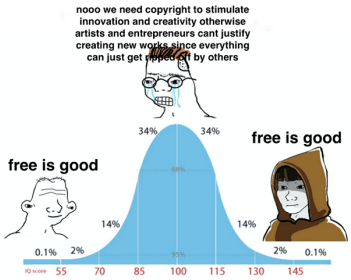
answered with: “You are Anna!” It’s a common name, like many others used in marketing and branding, meant to create a sense of connection with the user and, perhaps, a sense of identification. Reading the blog and observing the development of Anna’s Archive over the past few years, it becomes clear that, despite the project’s initial anonymity, which seemed focused exclusively on the technical aspects of resistance to censorship, Anna has gradually become more self-conscious, developing a personality and style that are unique.

For example, Anna does not have an “About” page that states what she is, who the founders are, or what their political views might be. Instead, she features a prominent FAQ page, which seems to be a literary genre in itself. It is structured as a long interview composed of questions submitted by users, the first being: “How can I help?” Most of the questions address the technical aspects of the platform, but by reading through the entire FAQ, one can gain a detailed understanding of how Anna works: downloading, donations, and source code. At the same time, certain quirks allow Anna’s personality to emerge: “I downloaded 1984 by George Orwell—will the police come to my door?”; “Who is Anna? You are Anna!”; and perhaps most striking of all, a book list at the end:

> What are your favorite books?
> Here are some books that carry special significance to the world of shadow libraries and digital preservation:
> Michele Boldrin, David K. Levine, *Against intellectual monopoly*;
> Stephenson, Neal, *Cryptonomicon*;
> Aaron Swartz, *The Boy Who Could Change the World : The Writings of Aaron Swartz*;
> Witt, Stephen R, *How Music Got Free : A Story of Obsession and Invention*;
> А. М. Прохоров, *Физическая энциклопедия. Ааронова – Длинные Том 1.*

Swartz, Boldrin, Levine, and Witt are among the most important figures in the Copyleft movement, and they outline a clear framework within which Anna operates and perhaps expresses a political statement (despite Anna stating in several places that she is not concerned with politics). Anna’s Archive team, much like in Boldrin and Levine’s work, criticizes the outdated copyright system, arguing that, due to the simplification of access to knowledge in the modern world, many aspects should be updated. For example, the 70-year copyright expiration period is significantly longer than that of patents, which expire after just 20 years. In other words, directly or indirectly, even while avoiding explicit political positioning in practice, Anna engages in the debate on copyright law, proposing reforms and challenging the notion that only profit derived from intellectual monopolies can produce quality work. Open-source projects such as Wikipedia, Linux, or Anna itself, serve as powerful evidence that the opposite may well be true.

Besides the discussion on copyright, Anna's radicalism emerges from a dismissal of and distrust in institutions regarding their broader role in preserving knowledge and sharing it with future generations:



Humanity largely entrusts corporations like academic publishers, streaming services, and social media companies with this heritage, and they have often not proven to be great stewards.

There are some institutions that do a good job archiving as much as they can, but they are bound by the law. As pirates, we are in a unique position to archive collections that they cannot touch, because of copyright enforcement or other restrictions. We can also mirror collections many times over, across the world, thereby increasing the chances of proper preservation (https://annas-archive.org/blog/blog-how-to-become-a-pirate-archivist.html).

Anna's simplicity and focus on the technical aspects of preservation force us to reconsider the true effectiveness and usefulness of conventional institutions.

The culture of mistrust towards corporations, institutions, governments, and laws reflects a broader crisis that perhaps began with the financial collapse of 2008 and the Occupy Wall Street movements. Mistrust of financial institutions quickly led to skepticism about anything related to them: virtually everything, including academic institutions, especially in the Anglosphere. The same years that saw the rise of Library Genesis also witnessed the creation of the first cryptocurrencies, developed as a response to an unreliable financial system. The main idea behind shadow libraries, as well as cryptocurrencies, is to decentralize the control and distribution of currency or written information. Crucially, what makes the difference is the technology behind the distribution and preservation of data or currency, based on sparse networks and redundant registers rather than centralized systems. This was the driving force behind the development of blockchain technology, specifically designed to eliminate the need for a central institution, such as a bank, to issue and regulate currency. In both cases, cryptocurrency and Anna's Archive, there is an underlying mistrust of institutions, which appear as increasingly unstable. Recent global events, such as wars and sanctions that have crippled globalization, such as the exclusion of Russian banks from the SWIFT system, demonstrated how cryptocurrencies can be effective in bypassing state-imposed restrictions, ensuring anonymity, and enabling hundreds of thousands of refugees to move to foreign countries. Anna primarily uses cryptocurrencies for its transactions. Systemic anonymity and peer-to-peer/blockchain technology appear to be tools to counteract growing global authoritarianism.

Beyond the simplicity of Anna, there is a highly challenging and technically advanced system for storing data across different servers and making it accessible through various pathways to the library. However, this simple approach may also be the answer to the need for communication in a completely global and cross-cultural manner, handling publications that range from Chinese scientific books to Canadian poetry. The radical nature and success of Anna lies in its lightness and clarity in mediating between producers and consumers of knowledge. From this perspective, it may seem almost provocative and difficult to accept, as we, as users of

media content, have grown accustomed to progressively complex graphics and services in the multimedia world. Furthermore, the general approach of ephemeral curatorship is antithetical to the classical Western academic and cultural world, where large institutions create weighty and culturally framed systems for the storage and dissemination of knowledge.

Anna, underfunded, illegal in many countries, and lacking a board of directors or a scientific committee, forces us to question the effectiveness of such institutions in the modern world. How is it possible that Anna is able to provide access to real knowledge to far more people, in a far more equal and efficient way, than anyone else? Its very existence leads us to suspect that the superstructures of these institutions may not truly be interested in advancing human knowledge, but rather in managing it under the influence of other powers. At best, they are simply not efficient enough. Ultimately, Anna is separating knowledge from culture, favoring the former, as she plainly states:

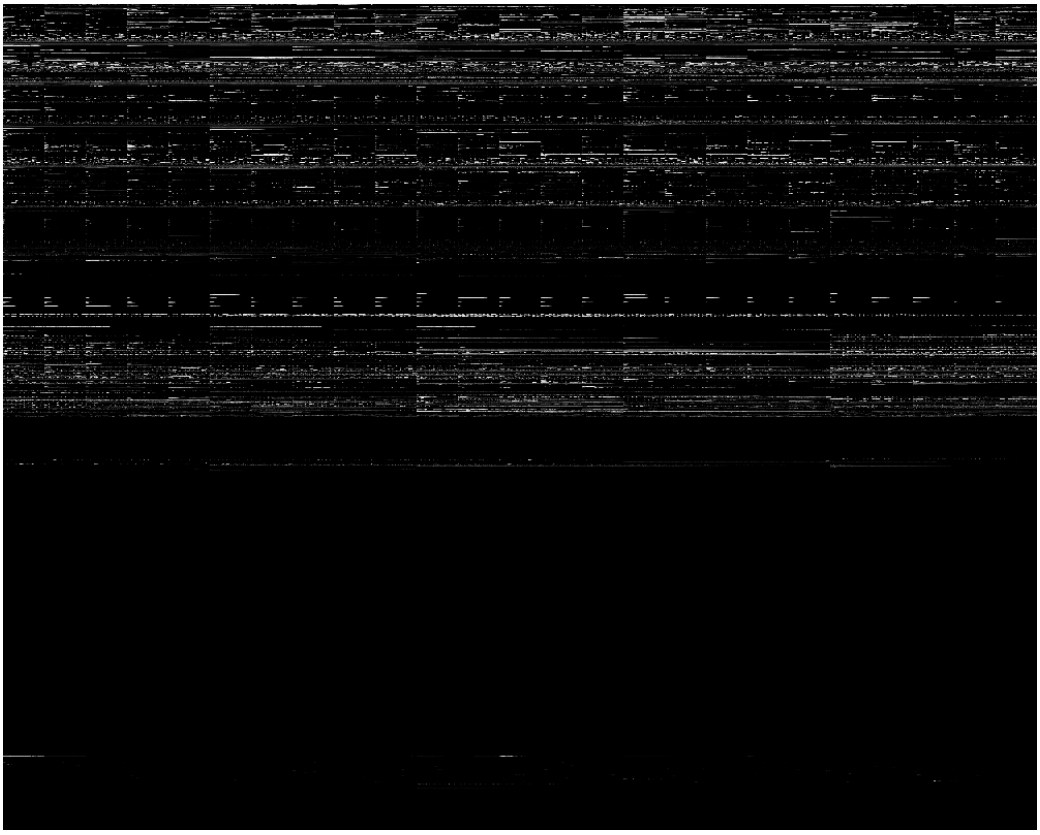
> Per megabyte of storage, written text stores the most information out of all media. While we care about both knowledge and culture, we do care more about the former (https://annas-archive.org/blog/critical-window.html).

Anna's speaking visualization

The radical essentiality of Anna naturally leads to the search for self-representation in a succinct way. This is no easy task. Anna is a library that aims to collect all the world's books, in a manner reminiscent of Borges' and Eco's works. Umberto Eco imagined a library where every book was identified by four numbers: one for the room, one for the wall, one for the shelf, and one for the position on the shelf. These four numbers would encode much of the information about the library itself, for example, the second number would reveal something about how many walls the rooms had. When we imagine archiving knowledge, mathematics and combination theory seems to be common features. Anna is no exception to this logic.

Recently, on Anna's blog, there was a prize competition for whoever could represent all published books in the best and most elegant way, a "To-Do List of Human Knowledge," as they called it. As with Eco's work, the main character in this representation is a number, but this time composed of 13 digits: the standard ISBN (International Standard Book Number), introduced in 1970, which identifies a specific edition of a book. Like Eco's numerical system, ISBNs are divided into parts: the first, 978, signals "this is an ISBN," followed by digits that indicate the group (a country or language area), the publisher, and the individual book. It is a practical system for managing the logistics of books as products, much like how other barcodes identify tomato sauce or toilet paper in a supermarket. ISBNs are simply numbers with a prefix, and they can cover up to 2 billion books. This is a massive number—many times greater than all books published to date—and a graphic representation of all ISBNs is the first step toward a general visualization of written human knowledge.

When describing the tasks of the competition, Anna published the simplest visualization: a large square where each pixel represents a possible ISBN, arranged in progressive order. Black pixels represent numbers that have not yet been assigned—white ones indicate published

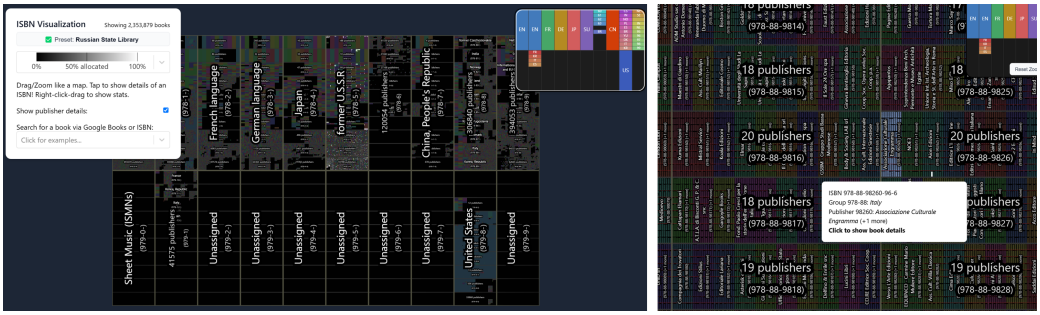


books. The image is already quite revealing we see a cryptic pattern composed of thousands of thin segments of varying lengths, ranging from the full width of the square to just a few pixels, unevenly distributed across the space. All books written by humans form a sparse constellation within the dark expanse of all possible books. These lines and segments resemble a complex combinatorial problem involving sets and subsets of a larger set, yet they are the result of human administration, a combination of regularity and randomness at multiple scales. ISBNs are assigned in large blocks, sometimes as many as 100 million, to different countries, which then determines how to distribute them. As a result, we can see the constellation breaking at regular intervals, with those intervals further subdivided in different ways—sometimes regularly, sometimes not, sometimes on a larger scale, sometimes on a smaller one—creating fractal-like structures.

In this image, humanity's library is not a neat row of books filling shelf after shelf, as we might see in a physical library or in the imagined libraries of Eco and Borges. Rather, it is a chaotic and mostly empty space, whose vast void is only sporadically interrupted by the presence of books. The main problem with this representation is that it arranges the space in lines,

causing individual clusters of books, such as, for example, all Brazilian literature, to be compressed into a horizontal segment just one pixel high. To address this, one of the winning competition entries attempts to arrange the ISBN sequence in a way that allows numerically adjacent entries to form a contiguous area or patch in two dimensions. To achieve this, it uses a space-filling curve, a special mathematical construct derived from Peano's work on space tessellation, originally developed to address the problem of the measurability of infinity.

The line snakes around the plane, outlining different groups in patches that resemble a geographical map, with winding fractal contours like those of an idealised coastline. Each coloured pixel inside the shapes represents a published book. Different countries, progressively smaller in size, are nested within one another, and can continue indefinitely, as long as there remains available space. Shapes are complex and unique, offering a visual idea of the number of books published in each language. English is by far the largest. The previously sparse constellation has now condensed in a singular geography of human knowledge, not based on traditional categorisation (such as the Dewey Decimal Classification in classic libraries), but instead according to an arbitrary set of rules and their mathematical visualisation. As in the previous image, these continents of books are more similar to archipelagos immersed in a dark sea. But now we can begin to compare them and see, as in a giant Risiko-style map, which continents are more dominant in the book production. Unsurprisingly, they don't differ significantly from their geopolitical counterparts.



The second representation, which won the prize, is, on the opposite, much more orderly, more similar to Borges' and Eco's imaginary libraries. Here books are arranged on regular virtual shelves, placed in regular cases, grouped into regular clusters, and so on, forming a progression of increasingly larger subsets and sets. In this case, the groups do not form interlocking shapes, but are instead a uniform row of rectangles with varying levels of infill. The previous fractal geography of human knowledge has been replaced by a cold-hearted, but more reassuring table, tidy enough to find whatever we are looking for, yet disturbing in the rigidity of the grid. It evokes an infinite Hippodamian city, where the plots are slowly but inevitably being developed, until the last one is eventually claimed.

These are just the initial attempts at what appears to be an overly complex exercise in data visualization, but Anna's Archive is one of the few libraries trying to achieve it. These efforts to represent an understated "To-Do-List" are raising the fundamental question of what constitutes a library or an archive today. Representing all of humanity's books makes human knowledge, at first glance, seem manageable, accessible, and controllable. However, upon delving into the details of these representations, we are quickly overwhelmed by the density of their internal structure and the intricate, irregular recursivity within them.

As libraries become larger, the method of their representation and the structure of their archive becomes the library itself. It is the combinatorics of its elements that makes it understandable and navigable, as envisioned by Eco and Borges. And perhaps Anna is more bizarre and complex than they would have ever imagined for a universal library. Yet, as an utopian library for the future, it brings their visions to life.

Anna and Artificial Intelligence

Since the universal library, in both its representation and structure, seems infinite and overwhelming to the human mind, it is nevertheless finite in number, meaning it can be scanned and processed with the help of large language models (LLMs). Anna realized quite early on that her archive would become an extremely valuable resource for the training of LLMs:

> This is a short blog post. We're looking for some company or institution to help us with OCR and text extraction for a massive collection we acquired, in exchange for exclusive early access. After the embargo period, we will of course release the entire collection.

> High-quality academic text is extremely useful for training of LLMs. While our collection is Chinese, this should be even useful for training English LLMs: models seem to encode concepts and knowledge regardless of the source language (https://annas-archive.org/blog/duxiu-exclusive.html).

Nowadays, artificial intelligence has become a strategic asset, used for data analysis, security, and military applications. We are witnessing an arms race in this field, and language models are particularly useful for extracting information from archives. In the not-so-distant future, any user will be able to ask an AI to retrieve, sort, and summarize all relevant information on a specific topic by searching through the entire world's library. This would be far more reliable and precise than current models trained to search the internet. And the more the model is trained to read texts printed since the 1400s, the better it will become at doing so, eventually being able to independently collect information and answer questions.

Current copyright rules are not only unfit for this new scenario, but even possibly dangerous, as Anna states:

> Me and my team are ideologues. We believe that preserving and hosting these files is morally right. Libraries around the world are seeing funding cuts, and we can't trust humanity's heritage to corporations either.
> Then came AI. Virtually all major companies building LLMs contacted us to train on our data. Most (but not all!) US-based companies reconsidered once they realized the illegal nature of our work. By contrast, Chinese firms have enthusiastically embraced our collection, apparently untroubled by its legality. This is notable given China's role as a signatory to nearly all major international copyright treaties (https://annas-archive.org/blog/ai-copyright.html).

The history of shadow libraries is likely to change soon due to the rise of Large Language Models (LLMs). One of the main challenges in training artificial intelligence is having access to high-quality, reliable data. This is precisely what shadow libraries offer: massive, well-organized collections of trustworthy material. If some corporations gain free access to these valuable archives while others remain restricted by copyright law, the former will have a significant advantage, as they could train their models using books written by experts, while the latter would be limited to smaller archives or unreliable, inconsistent sources from the web and social media. This is exactly what is happening now with Chinese companies training their AI models using Anna's Archive. And it may signal the beginning of the end for open libraries and Open Access more broadly. In the future, countries and corporations might stop publishing content openly to prevent giving their competitors an edge. The Universal Library could eventually become the Universal Machine Library, and shadow libraries might restrict full access to limit exploitation by artificial intelligence.

Conclusion

While at the beginning of the 20th century the question concerned the state of a work of art in the age of mechanical reproduction, now, in the 21st century, we are facing a reaction to that era—marked by the cessation of reproducibility. A similar process is unfolding in the dissemination of knowledge. Whereas there were once efforts to promote mass literacy and

education, we are now witnessing a tendency to restrict the distribution of knowledge – often under the pretext of security concerns or the protection of individual authors' rights. From Anna's perspective, we might speak of a kind of intellectual reactionism, a shift toward the elitization of knowledge, which is becoming increasingly evident. What does this symptom reveal? Will the trends shaping the future of science become more apparent in time? Does it suggest that mass science is no longer needed for some reason? But if so, what then does a library become? And what is Anna's role in this context?

Against this background, Anna's Archive should appear helpless. Instead, radically simple, incredibly intricate in its functioning, mirrored in thousands of copies, ungraspable, openly anticultural and anti-academic, it responds to these developments, attempting to spark a dialogue about the copyright system. This dialogue is nearly impossible, however, due to her explicit "shadowiness." Yet the real response to the emerging forms of global authoritarianism is not found in her words or blocked blogs, but in the project itself. Anna is a massive act of protest about what is happening. She is a true underground of knowledge, a utopia resisting the threats of contemporary dystopias. But even in the construction of an ideal utopian library of the future, old questions inevitably arise: even if you copy everything, where do you start? What is most important, books? Articles? Novels? Scientific papers? Works of natural philosophy or ancient literature? Even a new kind of library cannot escape the longstanding issue of the hierarchy of genres, even if that hierarchy is expressed solely through ISBN codes. Anna, in her radical simplicity, seeks to confront this in the most rational way possible, inventing new methods for overseeing, that is, representing, all of humanity's books. It is the project that has, so far, come closest to both the creation and visualization of a truly universal library, not a writer's speculation, nor a cultural ornament, but a highly technical system of preservation and distribution, designed to function seamlessly across global cultures.

Perhaps Anna is not a cultural phenomenon, but rather just a pirate site, another secret tool for academics and students. Regardless, even if she is merely a torrent site, are all these blockings necessary? Is it truly essential to subject knowledge databases, most of which are created with publicly funded research, to sanctions? Will these actions have any effect? And will they have an unexpected one? As academics, we may selfishly hope that Anna will persist for some time in the struggle for the right to free knowledge, and that it will reach at least 10% of all downloaded books. And then, perhaps, we will be able to get a little closer to answering the questions: What is Anna? Who is Anna? Perhaps the best clue comes again from Anna's peculiar FAQ bibliography, specifically from the last, and curiously unusual book—an old Soviet physics encyclopedia (Prohorov 1988), which begins alphabetically with the Aharonov-Bohm effect, bearing the same name as Aaron Schwartz.

> Quantum mechanic effect, characterised by the influence of an external electromagnetic field, concentrated in a region that is not accessible, on the quantum state of a charged particle (Prohorov 1988, 7).

A foundational effect of modern quantum physics, the Aharonov-Bohm effect, states that particles can be affected by a field even when the field is enclosed, due to an invisible coupling between the wave function and the electromagnetic potential. This concept mirrors Aaron Schwartz's activism: even when confined and repressed, his influence can affect distant minds, in charged particles or free spirits, transcending the commonly accepted principles of locality. Like the quantum effect, his actions can resonate globally, causing a "phase shift", the Anna effect.

Bibliographical References

Adams 2025
C. Adams, *Vanishing Culture: When Preservation Meets Social Media,* Internet Archive Blogs, 9 April 2025.

Aiguo (2009).
L. Aiguo, *Digitizing Chinese Books: A Case Study of the SuperStar DuXiu Scholar Search Engine,* "Journal of Academic Librarianship", 2009, 35, 277-281.

Anna 2022-2025 I
Anna's Blog.

Anna 2022-2025 II
Anna's Archive blog on Reddit.

Bodó 2018
B. Bodó, *The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library*, in Karaganis et al. 2018, 25-51.

Boldrin 2008
M. Boldrin, D. K. Levine, *Against Intellectual Monopoly*, Cambridge, 2008.

Correa, Laverde-Rojas, Tejada, et al. 2022
J.C. Correa, H. Laverde-Rojas, J. Tejada, et al., *The Sci-Hub effect on papers' citations,* "Scientometrics" 2022, 127, 99–126.

Karaganis et al. 2018
J. Karaganis (ed.), *Shadow Libraries: Access to Knowledge in Global Higher Education*, Boston 2018.

Pedersoli, Toson 2020
A. Pedersoli, C. Toson, *Onde libere e rock 'n' roll. La rivoluzione delle emittenti* offshore, "La Rivista di Engramma" n. 174, luglio/agosto 2020, 25-76.

Prohorov 1988
A.M. Prohorov, *Fizicheskaya entsiklopediya. AAronova – Dlinnye tom 1*, Moscow 1988.

Ruhenstroth 2014
M. Ruhenstroth, Schattenbibliotheken: Piraterie oder Notwendigkeit?, interview at Balázs Bodó, iRights info, 10 October 2014.

Rumfitt 2022
A. Rumfitt, *In defence of Z-Library and book piracy*, dazeddigital.com, November 2022.

Slum 2025
SLUM: The Shadow Library Uptime Monitor.

Stephenson 1999
N. Stephenson, *Cryptonomicon*, New York, 1999.

Swartz 2015,
Aaron Swartz, *The Boy Who Could Change the World: The Writings of Aaron Swartz*, New York, 2015.

Van der Sar 2025 I
E. Van der Sar, *Pirate Libraries Are Forbidden Fruit for AI Companies. But at What Cost?*, torrentfreak.com, 31 January 2025.

Van der Sar 2025 II
E. Van der Sar, *Anna's Archive Urges AI Copyright Overhaul to Protect National Security*, torrentfreak.com, 1 February 2025.

Wikipedia 2025
Wikipedia page for Shadow library, accessed April 2025.

Witt 2015
S. R. Witt, *How Music Got Free: A Story of Obsession and Invention*, New York, 2015.

Abstract

This paper attempts to interpret Anna' Archive in the context of the so called "Shadow Libraries", and its effects on the global circulation of knowledge and the debate over copiright. Anna is the last of a long series of online repositories of shared written material that originated in the early years of the World Wide Web, especially in post-Soviet academic environments. Anna has a unique approach to digital preservation, based on a radical refusal of any law or insitution that limits the possibility of access and conservation of knowledge, encouraging the act of copying or mirroring her data. Anna's style has developed from the beginning of the project, and has slowly developed a personality and a sophisticated work procedure. At the same time, Anna seems to seek a simple and radical approach, in order to act at a truly global scale. Anna is perhaps the only libary in the world that aspires to be universal, and is attempting to catalog and visualise the entire humanity's written collection in innovative ways, that remind of the fictional libraries immagined by Borges and Eco, and questions us on what should be the real nature of a library in the future.

keywords | Anna's Archive; universal library; shadow library; Intellectual property; Knowledge representation; AI training datasets.

la rivista di **engramma**
marzo **2025**
222 • Copyleft & internauti pirati

Editoriale
Filippo Perfetti, Giulia Zanon
“©Tutti i diritti riservati”. Proprio tutti?
Peppe Nanni
Engramma Open Access. Aperta origine
Redazione di Engramma

Invenzione
Riprodurre il patrimonio culturale
Mirco Modolo
Per un’iconografia della noia
Lorenzo Gigante
A Tale of Two Misfits
Maurizia Paolucci
Una lettura di: Yan Thomas, Il valore delle cose, Quodlibet [2002] 2022
Bernardo Prieto

Chiusura
No logo?
Giorgiomaria Cornelio
Il potere segreto. Wikileaks e la digitalizzazione dell’informazione
a cura di Alessandro Visca
Museums and the Enclosure of the Public Domain in the Digital Age
Douglas McCarthy
Piracy Shield, diritto d’autore e monopoli intellettuali: il caso italiano
Barbara Pasa
Note su alcune aperture giurisprudenziali al principio del “libero utilizzo” e sugli effetti nelle culture del riuso
Alessia Brandoni

Apertura
Copyright and Copyleft nell’era dell’intelligenza artificiale
Alessandro Del Ninno
Contro il copyright
Francesco D’Isa
Anna, the Universal Library
Elizaveta Kozina, Christian Toson
Per una liberazione dalla retorica dell’archivio
Alessandro Gagliardo con Filippo Perfetti e Giulia Zanon

Dalla cambusa
“Corrispondenza con un editore”
a cura di Filippo Perfetti
Guerilla Open Access Manifesto
Aaron Swartz, a cura di Giulia Zanon
Elogio de la piratería
Mauricio Souza Crespo, a cura di Filippo Perfetti
10 tesi sull’archivio
gruppo pad.ma, edizione italiana a cura della redazione di Inannunciato