UNIVERSITY OF AMSTERDAM

UvA-DARE (Digital Academic Repository)

The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library

Bodó, B.

Publication date
2018
Document Version
Final published version
Published in
Shadow Libraries

Link to publication

Citation for published version (APA):
Bodó, B. (2018). The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library. In J. Karaganis (Ed.), *Shadow Libraries: Access to Knowledge in Global Higher Education* (pp. 25-51). MIT Press. https://www.ivir.nl/publicaties/download/library_genesis.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:05 Sept 2026

2 The Genesis of Library Genesis: The Birth of a Global Scholarly Shadow Library

Balázs Bodó

Here's what I see as a consequence of free educational book distribution: within decades, generations of people everywhere in the world will grow up with access to the best scientific texts of all time. […] [T]he quality and accessibility of education to the poor will grow dramatically too. Frankly, I see this as the only way to naturally improve mankind: we need to make all the information available to them at any time.

—Anonymous administrator of the Russian shadow library site Library Genesis (LG), explaining its raison d'être

(Pirate) Libraries on the Internet

Digital librarianship—the digitization, collection, and cataloguing of texts—was one of the earliest uses of networked computers. By most accounts, the first digital library was Project Gutenberg, which began making public domain works available in 1971 via the Arpanet, the predecessor of the Internet. As computing and network technologies improved in the 1980s and 1990s, the technical obstacles and cost of building digital libraries declined rapidly. The dream of a universal library (Battles 2004; Borges 1998; Bush 1945; Rieusset-Lemarié 1997) began to seem very real. Legal obstacles were another matter. As projects became larger and more visible, they became more vulnerable to copyright challenges in the poorly charted areas around digitization, archiving, and fair use. Some projects responded by moving texts into closed, "dark" collections, maintained offline.[1] Others worked to assert and clarify rights to digitization and online distribution, prompting a flurry of lawsuits from publishers and authors' groups.

Major lines of conflict passed through lawsuits against big players like Google Books and the Hathi Trust, which represented a coalition of universities. Provisionally and only under U.S. law, these cases settled important questions about fair use in digitization projects and the handling of "orphan" works, for which the copyright holder

could not be identified. Other conflicts emerged around the scope of permissible use of copyrighted materials in educational contexts—particularly in lawsuits against university libraries and copy centers. Still others involved enforcement against projects that saw free digital libraries as ideological projects—as fundamental social goods.

The latter projects were generally small in scale, volunteer based, clandestine, and sometimes accidental in their origins—personal archives that grew into shared collections. The development of organizational and bibliographical infrastructure was a major challenge for such projects and a signal of larger ambitions. Some remained simple collections of texts bundled and exchanged via DVDs, torrents, or IRC channels online. Others acquired many of the attributes of libraries, including the crucial one: the catalog.

The catalog distinguishes an unstructured heap of computer files from a collectively managed and maintained collection of texts. For users, it has obvious utility for searching and browsing the collection. But it is also the organizing framework for the community of "librarians" who preserve and nourish the collection. The significant academic shadow libraries of the past decade—Textz.org, a*.org, monoskop, Gigapedia (later known as Library.nu), and more recently LibGen and Sci-Hub—took shape and gained traction through cataloguing efforts. Most maintained a bifurcated structure, in which the catalog serves as a platform for searching, organizing, and community engagement, while the actual texts are hosted elsewhere. This was partly a matter of convenience but also safety, as the legal system struggled to draw distinctions between searching, indexing, hosting, and other online functions.

As with the major music file sharing services in the early 2000s, public catalogs made shadow libraries easier targets of law enforcement. All of these services have faced takedown threats and, in several cases, injunctions that targeted the catalog, the text repository, or both. Of these libraries, Gigapedia/Library.nu—was the largest at the turn of the 2010s. At its peak, it was several orders of magnitude bigger than any of its peers, offering access to nearly a million English-language documents. It was not just size that made Gigapedia unique. Unlike most sites, which specialized in literary works, Gigapedia had large collections drawn from a wide range of academic disciplines, especially the sciences. Compared to its peers, it also had a highly developed central database, which contained bibliographic details on the collection and also, significantly, on gaps in the collection, which informed a process of soliciting contributions from users. With scanner and copiers now ubiquitous, users responded to requests and fueled the rapid growth of the collection.

In general, the major academic publishers were wary of following the music and film industries into a game of enforcement "whack-a-mole" against file sharing sites,

pouring resources into lawsuits against services that that would be reconstituted quickly under new identities and in other jurisdictions. But such reticence was not universal. By 2010, the apparent size of the Gigapedia had convinced several publishers that it was too big a target to ignore. Led by John Wiley & Sons, a group of seventeen publishers was granted an injunction against the site (called by then Library.nu) and against iFile.it—the hosting site that stored most of Library.nu's content. Under the injunction, the Library.nu administrators closed the site. From the outside, it seemed that the collection had disappeared and the community around it dispersed (Liang 2012). But provisions for the next Library.nu were in place well before the shutdown, circulating primarily through networks of Russian academics and shadow librarians. For reasons that we will explore in this chapter, much of the history of the big digital shadow libraries passes through Russia and the story of Library Genesis.

Library Genesis

Library Genesis[2] (also known as LG or LibGen) is a shadow library started by Russian scientists around 2008 to consolidate the mostly Russian-language text collections circulating on the Russian-language Internet. In 2011, LibGen swallowed the much larger and broader Library.nu collection.[3] For the LibGen community, Library.nu was just another free-floating text archive, ready to be harvested and integrated into the rest of the collection. But with the closure of Library.nu, LibGen inherited the responsibility of serving a larger academic community beyond the boundaries of Russian-speaking academia. The whacking of the Gigapedia mole gave rise to a mole with a large family and a more sophisticated and resilient approach to collecting and sharing books.

As a shadow library and piratical content distribution service, LibGen has a unique modus operandi. Most such websites tend to exercise strict control over the content they make accessible and the infrastructure they build. LibGen's mission, in contrast, is to provide open access to the collection by making *itself* radically open. It collects free-floating scientific texts and other collections from the Internet and consolidates them (both content and metadata) into a single, open database. Although ordinary users can search the catalog and retrieve the texts, LibGen's main focus is the distribution of its own library infrastructure, including its source code, catalog, and terabyte-sized collection to anyone who wants to start his or her own library. In practical terms this means that anyone can freely take a copy of LibGen and start distributing text under his or her own terms. This openness has led to the creation of a lively ecosystem of shadow libraries around the core LibGen collection. The ability to mirror LibGen without restrictions enables these sites to target different audiences by combining the LibGen catalog with

books coming from other sources,[4] providing extra services,[5] or experimenting with different financing models.[6]

This two-layered structure enables LibGen to focus its limited resources on maintaining a high-quality scientific collection, while the mirror sites compete to best serve users, carry the costs of distribution, and act as lightning rods for lawsuits. The mirror sites deliver the LibGen collection to the public, and at the same time, increase the likelihood of its long-term survival.

The main mission of the LibGen community is the development of the collection. Its democratic approach to access is matched by an elitist approach to content. As characterized by one of its administrators (admins), these goals are to

- collect valuable science/technology/math/medical/humanities academic literature. That is, collect humanity's valuable knowledge in digital form. Avoid junky books. Ignore "bestsellers."
- build a community of people who share knowledge, improve quality of books, find good and valuable books, and correct errors.
- share the files freely, spreading the knowledge altruistically, not trying to make money, not charging money for knowledge.

LibGen's agenda is marked by deep aversion to a narrowly academic understanding of research and education, especially with regard to elite institutions that provide gated access to knowledge for their communities. Instead, LibGen's statement takes the autodidacticism necessary to education in many parts of the world and reimagines it as a liberatory agenda—a future of self-learning communities based on universal access to knowledge. The LibGen admin further describes site priorities:

> The overwhelming arrogance of university staff will gradually be suppressed for a larger flow of exceptionally educated people without special degrees acquired (I am proudly the case, that's why I'm saying this, it's not a fantasy). […]
>
> The target groups for LibGen are poors: Africa, India, Pakistan, Iran, Iraq, China, Russia and post-USSR etc., and on a separate note, people who do not belong to academia. If you are not at a university, you can't access anything or at least your access will be so much troubled that you won't be able to progress at all.

It is easy to see parallels between LibGen and the agenda of someone like Aaron Swartz in the United States, whose Guerilla Open Access Manifesto touched on many of the same themes in 2008. (Swartz committed suicide in 2013 while under investigation for the unauthorized downloading of large parts of the JSTOR catalog of academic articles). As the technologically possible library surpasses the modest reality and uneven distribution of actual libraries, this sense of relative deprivation can readily become a politics. As Swartz puts it:

> Those with access to these resources—students, librarians, scientists—you have been given a privilege. You get to feed at this banquet of knowledge while the rest of the world is locked out. But

you need not—indeed, morally, you cannot—keep this privilege for yourselves. You have a duty to share it with the world. And you have: trading passwords with colleagues, filling download requests for friends.

Meanwhile, those who have been locked out are not standing idly by. You have been sneaking through holes and climbing over fences, liberating the information locked up by the publishers and sharing them with your friends.

But all of this action goes on in the dark, hidden underground. It's called stealing or piracy, as if sharing a wealth of knowledge were the moral equivalent of plundering a ship and murdering its crew. But sharing isn't immoral—it's a moral imperative.

The LibGen and Swartz manifestos are remarkably similar documents. There are, however, enormous differences in the contexts in which these texts were born, put into practice, and took effect (Bodó 2016). The limited Russian "success" in building large online shadow libraries where Swartz and many other shadow libraries documented in this study failed is not accidental. The dissonance that Swartz charismatically embodied in the United States within a community of hackers and activists was, to a degree, a societal experience in Russia, passing through clandestine publication practices under Soviet rule, the economic ruin of the intelligentsia in the post-communist period, and the weak legal infrastructure for copyright (and other law) that allowed a pirate Internet to flourish.

The Communist Ideal of the Reading Nation

[T]he library of the Big Lubyanka was unique. In all probability it had been assembled out of confiscated private libraries. The bibliophiles who had collected those books had already rendered up their souls to God. But the main thing was that while State Security had been busy censoring and emasculating all the libraries of the nation for decades, it forgot to dig in its own bosom. Here, in its very den, one could read Zamyatin, Pilnyak, Panteleimon Romanov, and any volume at all of the complete works of Merezhkovsky. (Some people wisecracked that they allowed us to read forbidden books because they already regarded us as dead. But I myself think that the Lubyanka librarians hadn't the faintest concept of what they were giving us—they were simply lazy and ignorant.)

—Aleksandr Solzhenitsyn, The Gulag Archipelago *1918–1956* (1974)

Russian culture has a deep reverence for the printed word and, in many respects, the Soviet state only amplified it. The Leninist program of education created a mass readership for the first time in Russia, but at the same time closed off many of the conventional outlets for it. As Stelmakh (1993) observed: "Reading almost transplanted religion as a sacred activity: in the secularized socialist state, where the churches were closed, the free press stifled and schools and universities politicized, literature became

the unique source of moral truth for the population. Writers were considered teachers and prophets."

The Soviet Union was a reading culture until the end. In the last days of the USSR, a quarter of the adult population were considered active readers, and almost everyone else qualified as an occasional reader (Stelmakh 1993). Book prices were low and alternative forms of entertainment were scarce and relatively expensive, making reading one of the most attractive leisure activities.

The communist approach toward intellectual property protection reflected the ideal of the reading nation. The Soviet Union inherited a lax and isolationist copyright system from the tsarist Russia. Neither the tsarist state nor the Soviet state adhered to international copyright treaties, nor did they enter into bilateral treaties. Tsarist Russia's refusal to grant protection to foreign authors and translations had an economic rationale: Russian publishers would pay no royalties for foreign work. The Soviet regime added a strong ideological claim: granting exclusive ownership to authors hindered "the cultural development of the masses," and only served the private interests of authors and heirs. As Elst (2005, 658) states: "If copyright had an economic function, that was only as a right of remuneration for his contribution to the extension of the socialist art heritage. If copyright had a social role, this was not to protect the author from the economically stronger exploiter, but was one of the instruments to get the author involved in the great communist educational project."

The Soviet copyright system, even in its postrevolutionary phase, maintained two features that served as important vehicles for new publication. First was the "freedom of translation," which meant that translations could be published without rights holder authorization. This measure dismantled a significant barrier to access in a multicultural and multilingual empire. By the same token, the denial of protection to foreign authors and rights holders eased the import of foreign texts (after, of course the appropriate censorship review). According to Newcity (1980, 6), due to these instruments: "[S]oon after its founding, the Soviet Union became as well the world's leading literary pirate, not only publishing in translation the creations of its own citizens but also publishing large numbers of copies of the works of Western authors both in translation and in the original language."

Looking simply at the aggregate numbers of published books, the USSR had an impressive publishing industry on a scale appropriate to a reading nation. Between 1946 and 1970, more than one billion copies of over twenty-six thousand different works were published, all by foreign authors (Newcity 1978). This production accelerated rapidly in the 1970s. In 1976 alone, more than 1.7 billion copies of 84,304 titles were printed[7] (Friedberg, Watanabe, and Nakamoto 1984, fn 4).

Censorship

Of course, these impressive numbers did not reflect a healthy public sphere or a well-functioning print ecosystem. The book-based public sphere was both heavily censored and plagued by the growing dysfunctions of the Soviet, and later the post-Soviet, economy.

The totalitarian Soviet state had many instruments to control the circulation of literary and scientific works.[8] Some texts never entered official circulation at all. As Stelmakh (2001, 145) notes: "A particularly harsh prepublication censorship [affected] foreign literature, primarily in the humanities and socioeconomic disciplines. Books on politics, international relations, sociology, philosophy, cybernetics, semiotics, linguistics, and so on were hardly ever published."

Many "problematic" texts were put into limited circulation for the trustworthy few. As the resolution of the Central Committee of the Communist Party of June 4, 1959, stated: "Writings by bourgeois authors in the fields of philosophy, history, economics, diplomacy, and law […] are to be published in limited quantities after the excision from them of passages of no scholarly or practical interest. They are to be supplied with extensive introductions and detailed annotations" (quoted in Friedberg, Watanabe, and Nakamoto 1984).

The truncation and mutilation of texts were also frequent. Literary works and texts from humanities and social sciences were obvious subjects of censorship, but natural sciences and technical fields did not escape. Dewhirst and Farrell (1973, 127) reported: "In our film studios we received an American technical journal, something like *Cinema, Radio and Television*. I saw it on the chief engineer's desk and noticed that it had been reprinted in Moscow. Everything undesirable, including advertisements, had been removed, and only those technical articles with which the engineer could be trusted were retained. Everything else, even whole pages, was missing. This was done by a photo copying process, but the finished product appeared to be printed."

Mass cultural genres were also subject to censorship and control. Women's fiction, melodrama, comics, detective stories, and science fiction were completely missing or heavily underrepresented in the mass market. Instead, "a small group of officially approved authors […] were published in massive editions every year, [and] blocked readers' access to other literature. […] Soviet literature did not fit the formula of mass culture and was simply bad literature, but it was issued in huge print-runs" (Stelmakh 2001, 150).

Libraries were also important instruments of censorship. When not destroyed altogether, censored works ended up in the *spetskhrans,* limited access special collections

established for censored works. Besides obvious candidates such as anti-Soviet works and Western "bourgeois" publications, many scientific works ended up in these closed collections (Ryzhak 2005). Access to the *spetskhrans* was limited to those with special permits: "Only university educated readers were enrolled and only those holding positions of at least junior scientific workers were allowed to read the publications kept by the spetskhran" (Ryzhak 2005). In the last years of the USSR, the *spetskhran* of the Russian State Library—the largest of them, with more than one million items in the collection—had forty-three seats for its roughly forty-five hundred authorized readers. Yearly circulation was around two hundred thousand items, a figure that included "the history and literature of other countries, international relations, science of law, technical sciences and others" (Ryzhak 2005).

Librarians thus played a central role in the censorship machinery. They did more than guard the contents of limited-access collections and purge the freely accessible stocks according to the latest Communist Party directives. As the intermediaries between the readers and the closed stacks, their task was to carefully guide readers' interests and report on suspicious reading habits (Stelmakh 2001).

Access to works was limited by economic factors as well. Due to the lack of signals for demand and the bureaucratic limitations of the planned economy, shortages of even censor-approved texts were common, both on the market and in libraries. Access to foreign works was further limited when the USSR joined the UNESCO-backed Universal Copyright Convention (UCC) in 1973. Under the UCC, the USSR finally granted protection to foreign authors and put an end to the "freedom of translation" clause—the exemption in Soviet author rights law that permitted the translation of works without the authorization of the rights holder. Soviet officials feared that granting protection to foreign authors would result in an outflow of royalty payments to Western rights holders. As data shows, these fears proved valid. By 1976, the annual USSR trade deficit in publishing reached a million rubles (around $5.5 million in current USD) (Levin 1983, 157). This imbalance also raised the price of translated works to double that of Russian-authored books (158).

The Soviet and Post-Soviet Literary and Scientific Underground

Various practices and informal institutions evolved to address the problems of access. Black markets for books flourished: "In the 1970s and 1980s the black market was an active part of society. Buying books directly from other people was how 35 percent of Soviet adults acquired books for their own homes, and 68 percent of families living in major cities bought books only on the black market" (Stelmakh 2001, 146). Book

copying and hoarding also became widespread strategies for dealing with the shortages. One administrator of the LibGen shadow library has vivid, firsthand memories of these times:

> People hoarded books: complete works of Pushkin, Tolstoy or Chekhov. You could not buy such things. So you had the idea that it is very important to hoard books. High-quality literary fiction, high-quality science textbooks and monographs, even biographies of famous people (writers, scientists, composers, etc.) were difficult to buy. You could not, as far as I remember, just go to a bookstore and buy complete works of Chekhov. It was published once and sold out and that's it. Dostoyevsky used to be prohibited in the USSR, so that was even rarer. Lots of writers were prohibited, like Nabokov. Eventually Dostoyevsky was printed, but in very small numbers.
>
> And also there were scientists who wanted scientific books and also could not get them. Mathematics books, physics—very few books were published every year, you can't compare this with the market in the U.S. Russian translations of classical monographs in mathematics were difficult to find.
>
> So, in the USSR, everyone who had a good education shared the idea that hoarding books was very, very important, and did just that. If someone had free access to a Xerox machine, they were [x]eroxing everything in sight. A friend of mine had an entire room full of [x]eroxed books.

From the 1960s onward, the ever-growing clandestine samizdat networks challenged the censors and provided access to both classics and information on the current state of Soviet society. Reaching a readership of around two hundred thousand, these networks operated in a networked, bottom-up manner. Each node in a chain of distribution copied the texts it received, and distributed the copies. These nodes also carried information backward, toward the authors of the samizdat publications.

In the immediate post-Soviet turmoil, access to print culture did not get any easier. Censorship officially ended, but so too did much of the state funding for the publishing sector. Mass unemployment, falling wages, and the resulting loss of discretionary income further undercut the shift toward market-based publishing models. The funding of libraries also dwindled, limiting new acquisitions (Elst 2005, 299–300). Economic constraints, in short, took the place of political ones. But in the absence of political repression, self-organizing efforts to address these constraints acquired greater scope of action. Slowly, the informal sphere began to deliver alternative modes of access to otherwise hard-to-get literary and scientific works.

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. The prominent place of Russian pirate libraries in the larger informal media economy—and of Russian piracy of music,

film, and other copyrighted work more generally—cannot be understood outside this history.

The Emergence of Do-It-Yourself Digital Libraries in RuNet

The copying of censored and uncensored works (by hand, typewriters, photocopiers or—later—computers), the hoarding of copied texts, the buying and selling of books on the black market, and the informal, peer-to-peer distribution of samizdat material were part of the everyday experience of educated Soviet and post-Soviet readers. The building and maintenance of individual collections and participation in the informal networks of exchange offered a sense of political, economic, and cultural agency—especially as the public institutions that supported the core professions of the intelligentsia fell into sustained economic crisis.

Digital technologies were integrated into these practices as soon as they appeared. As one shadow library administrator remembers:

> From late 1970s, when first computers became used in the USSR and printers became available, people started to print forbidden books, or just books that were difficult to find, not necessarily forbidden. I have seen myself a print-out on a mainframe computer of a science fiction novel, printed in all caps! Samizdat was printed on typewriters, xeroxed, printed abroad and xeroxed, or printed on computers. Only paper circulated. Files could not circulate until people started to have PCs at home. As late as 1992 most people did not have a PC at home. So the only reason to type a long text into a computer was to print it on paper.

People who worked in academic and research institutions were well positioned to support these informal practices: they had access to computers, and many had access to the materials locked up in the *spetskhrans*. Many also had the time and professional motivations to collect and share otherwise inaccessible texts. The core of current digital collections was created in this late-Soviet/early post-Soviet period by such professionals. Their home academic and scientific institutions continued to play an important role in the development of digital text collections well into the era of home computing and the Internet.

Digitized texts first circulated in printouts and later on optical/magnetic storage media and the early Internet. The first platform for digital text sharing was the Russian Fidonet, a network of bulletin board systems similar to Usenet, which enabled the mass distribution of plain text files. These bulletin board systems (BBSs) connected fans around emerging collections of shared texts, such as the Holy Spirit BBS's "SU.SF & F.FANDOM" group, whose main focus was Soviet-Russian science fiction and fantasy literature. As one of the shadow librarians described their experience in the early 1990s:

Fidonet collected a large number of plaintext files in literature / fiction, mostly in Russian, of course. Fidonet was almost all typed in by hand. […] Maybe several thousand of the most important books, novels that "everyone must read" and such stuff. People typed in poetry, smaller prose pieces. I have myself read a sci-fi novel on a mainframe, which was obviously typed in. This novel was by Strugatski brothers. It was not prohibited or dissident literature, but just impossible to buy in the stores. These were culturally important, cult novels, so people typed them in. […] At this point it became clear that there was a lot of value in having a plain-text file, and the most popular novels were first digitized in this way.

The next stage in text digitization started around 1994. By that time, growing numbers of people had access to computers, scanning peripherals, and OCR (text recognizing) software. Household Internet and PC penetration, while extremely low overall in the 1990s, (0.1 percent of the population had Internet access in 1994, growing to 8.3 percent by 2003), began to make inroads in educational and scientific institutions and among Moscow and St. Petersburg elites, who were often the critical players in these networks. As access to technologies increased, a much wider array of people began to digitize their favorite texts. These collections began to circulate, first via CD-ROMs and later on the Internet.

Maxim Moshkov and lib.ru

One such collection belonged to Maxim Moshkov, who published his library under the name lib.ru in 1994. Moshkov was a graduate of the Moscow State University Department of Mechanics and Mathematics, which (as we'll see later) played a large role in the digitization of scientific works. After graduation, he worked for the Scientific Research Institute of System Development—a computer science institute associated with the Russian Academy of Sciences. He describes the early days of his collection as follows:

I began to collect electronic texts in 1990, on a desktop computer. When I got on the Internet in 1994, I found lots of sites with texts. It was like a dream came true: there they were, all the books I desired. But these collections were in a dreadful state! Incompatible formats, different encodings, missing content. I had to spend hours scouring the different sites and directories to find something.

As a result, I decided to convert all the different file-formats into a single one, index the titles of the books and put them in thematic directories. I organized the files on my work computer. I was the main user of my collection. I perfected its structure, made a simple, fast and convenient search interface and developed many other useful functions and put it all on the Internet. Soon, people got into the habit of visiting the site. […]

For about two years I scoured the [I]nternet. I sought out and pulled texts from the network, which were lying there freely accessible. Slowly the library grew, and the audience increased with it. People started to send books to me, because they were easier to read in my collection. And the

time came when I stopped surfing the [I]nternet for books: regular readers now send me books. Day after day I get about 100 emails, and 10–30 of them contain books. So many books came in that I did not have time to process them. Authors, translators, and publishers also started to send texts. They all needed the library. (Moshkov 1999)

In the second half of the 1990s, the Russian Internet—RuNet—was awash in book digitization projects. With the advent of scanners, OCR technology, and the Internet, the work of digitization had eased considerably. Texts migrated from print to digital and sometimes back to print again. They circulated through different collections, which, in turn, merged, fell apart, and reformed. Digital libraries with the mission to collect and consolidate these free-floating texts sprung up by the dozens.

Such digital librarianship was the antithesis of official Soviet book culture: it was free, bottom-up, democratic, and uncensored. It also offered a partial remedy to problems created by the post-Soviet collapse of the economy: the impoverishment of libraries, readers, and publishers. In this context, book digitization and collecting also offered a sense of political, economic, and cultural agency, with parallels to the copying and distribution of texts in Soviet times. The capacity to scale up these practices coincided with the moment when anti-totalitarian social sentiments were the strongest, and economic needs most dire.

This unprecedented bloom of digital librarianship was the result of the superimposition of multiple waves of technological, political, economic, and social transformation. "Maksim Moshkov's Library" was ground zero for this convergence and soon became a central point of exchange for the community engaged in text digitization and collection: One shadow librarian recalled this period as follows: "[At the outset] there were just a couple of people who started scanning books in large quantities. Literally hundreds of books. Others started proofreading, etc. There was a huge hole in the market for books. Science fiction, adventure, crime fiction—all of this was hugely in demand. Lib.ru was a large part of the response, and was filled with the books that people most desired and valued."

For years, lib.ru integrated as much as it could of the different digital libraries that flourished in the RuNet—preserving, in the process, many of the smaller, short-lived libraries.

This process of collection slowed in the early 2000s. By that time, lib.ru had all of the classics, resulting in a decrease in the inflow of new material. By the same token, the Russian book market was finally starting to offer works aimed at the mainstream, resulting in an abundance of romances, astrology, crime fiction, and other popular genres. These works started to appear in, and would soon flood, lib.ru. Many contributors, including Moshkov, were concerned that such ephemera would dilute the

original library. And so they began to disaggregate the collection. Self-published literature, "user-generated content," and fan fiction were separated into the aptly named samizdat section of lib.ru (http://samlib.ru/), which housed original texts submitted by readers. Popular fiction—"low-brow literature"—was split off. Sites specializing in those genres quickly formed their own ecosystem. Librusec, the first of its kind, now charges a monthly fee to provide access to the collection. The Flibusta community split off from Librusec the same way that Librusec split off from lib.ru, to provide free and unrestricted access to a similar collection. Finally, some in the community felt the need to focus their efforts on a separate collection of scientific works. This became the Kolkhoz collection.

Toward a Million-Book Scientific Library

A *kolkhoz* (Russian: колхо́з) was a type of collective farm that emerged in the early Soviet period. In those early days, it was a self-governing, community-owned collaborative enterprise, with many of the features of a commons. For the Russian digital librarians, these historical resonances were intentional. As the LibGen administrator described:

> The [K]olkhoz group was initially a community that scanned and processed scientific materials: books and, occasionally, articles. The ethos was free sharing. Academic institutes in Russia were in dire need of scientific texts; they xeroxed and scanned whatever they could. Usually, the files were then stored on the institute's FTP site and could be downloaded freely. There were at least three major research institutes that did this back in early 2000s, unconnected to each other in any way, located in various faraway parts of Russia. Most of these scans were appropriated by the [K]olkhoz group and processed into DJVU.[9]
>
> The sources of files for [K]olkhoz were, initially, several collections from academic institutes, downloaded whenever the FTP servers were open for anonymous access. In one case, this included one of the institutes of the Chinese Academy of Sciences, but mostly they came from Russian academic institutes. At that time [around 2002], there were also several commercialized collections of scanned books on sale in Russia. Mostly, these were college-level textbooks on math and physics. These files were also all copied to [K]olkhoz and processed into DJVU. The focus was on collecting the most important science textbooks and monographs of all time, in all fields of natural science.
>
> There was never any commercial support. The [K]olkhoz group never had a web site with a database, unlike most projects today. They had an FTP server with files, and the access to FTP was given by PM [one of the administrators] in a forum. This server was privately supported by one of the members—an academic researcher, like most [K]olkhoz members. The files were distributed directly by burning files on writable DVDs and giving them away. Later, FTP access was closed to the public and only a temporary file-swapping FTP server remained. Today the [K]olkhoz DVD releases are mostly spread via torrents.

The Kolkhoz collection amassed around fifty thousand documents. The *mexmat* collection of the Moscow State University Department of Mechanics and Mathematics (Moshkov's alma mater) was around the same size. The "world of books" (*mirknig*) collection had around thirty thousand files, and there were roughly a dozen other smaller archives with approximately ten thousand files in their respective collections.

The Kolkhoz group dominated the science-minded e-book community in Russia well into the late 2000s. Kolkhoz, however, suffered from the same problems as the early Fidonet-based text collections. Since it was distributed on DVDs, via FTP servers and later on torrents, it was hard to search, it lacked a proper catalog, and it was prone to fragmentation. Parallel solutions soon emerged. Around 2006–2007, the early Gigapedia copied the English books from Kolkhoz, set up a catalog, and soon became the most influential pirate library in the English-speaking Internet.

Similar cataloguing efforts soon emerged elsewhere. In 2007, someone on rutracker.ru, a Russian file sharing site, posted torrent links to ninety-one DVDs containing science and technology titles aggregated from various Russian sources, including Kolkhoz. This massive collection had no categorization or particular order. But it soon attracted a librarian: a user of the forum started the laborious task of organizing the texts into a usable, searchable format—first filtering duplicates and organizing existing metadata into an Excel spreadsheet, and later moving to a more open, web-based database. And thus Library Genesis was born.

LibGen inherited more than just books from Kolkhoz and Moshkov's lib.ru. It inherited their elitism with regard to canonical texts, and their understanding of librarianship as a community effort. Like the earlier sites, LibGen's collections are expanded by user submissions. Like the other sites, the number of submissions grew rapidly as the site's visibility, reputation, and trustworthiness were established, and like the others, this growth trailed off as the collection of canonical literature grew more complete. As the LibGen administrator explained:

> The number of mankind's useful books is roughly what we already have. So growth is defined by newly scanned or issued books. Also, the quality of the collection is represented not by the number of books but by the amount of knowledge it contains. LibGen does not need to grow further and I am not the only one among us who thinks so. […]
>
> We have absolutely no idea who sends books in. It is practically impossible to know, because there are a million books. We gather huge collections which eliminate any traces of the original uploaders.
>
> My expectation is that new arrivals will dry up. Not completely, as I described above. Some books will always be scanned or rescanned (it nowadays happens quite surprisingly often) and the overall process of digitization cannot and should not be stopped. It is also hard to say when the slowdown will occur: I expected it about a year ago, but then Library.nu got shut down and things

changed dramatically in many respects. Now we are "in charge" (we had been the largest anyways, just now everyone thinks we are in charge) and there has been a temporary rise in the book inflow. At the moment, relatively small or previously unseen collections are being integrated into LibGen. Perhaps in a year it will saturate.

However, intuition is not a good guide. There are dynamic processes responsible for [e-book] availability. If publishers massively digitize old books, they'll obviously be harvested and that will change the whole picture."

The ambitions of LibGen's administrators to create a universal library are limited, at least in terms of scope. It is not intended to contain everything. Its boundaries are created in dialogue with the community, measured by the act of actively digitizing and sharing books. Yet the size of this community is carefully limited. The administrators identified Gigapedia's visibility as the main contributor to its downfall and they wish to avoid that trap. On the one hand, as one admin stated: "Our policy, which I control as strictly as I can, is to avoid fame. Gigapedia's policy was to gain as much fame as possible. Books should be available to you, if you need them. But let the rest of the world stay in its equilibrium. We are taking great care to hide ourselves and it pays off."

On the other hand, LibGen's administrators understand that hiding limits the likelihood that scholars in need can find them. Their solution to this dilemma is to open source their collection and thereby allow others to create better publicized services that interface with the public. They let others run the risk of getting famous.

Copyright and "Copynorms" in Russian Pirate Librarianship

Library Genesis serves as a source archive for around a half-dozen freely accessible pirate libraries on the Internet. The catalog database is downloadable, the content is downloadable, even the server code is downloadable. No passwords are required to download and there are no gatekeepers. There are no obstacles to setting up a similar library with a wider catalog, an improved user interface and better services, a different audience or, in fact, a different business model.

This arrangement creates a two-layered community. The core group of LibGen admins maintains the current service, while a loose and ever-changing network of mirror sites build on the LibGen infrastructure. As the admins explained:

The unspoken agreement is that the mirrors support our ideas. Otherwise we simply do not interact with them. If the mirrors support this, they appear in the discussions, on the Web etc. in a positive context. This is again about building a reputation: if they are reliable, we help with what we can, otherwise they should prove [to] the World they are good on their own. We do not request anything from them. They are free to do anything they like. But if they do what we do not agree with, it'll be taken into account in future relations. If you think for a while, there is no

other democratic way of regulation: everyone expresses his own views and if they conform with ours, we support them. If the ideology does not match, it breaks down.

Forum posts asking for donations suggest that funding for LibGen comes from their own personal resources as well as occasional donations when there is a need to buy or rent equipment or services: "[W]e've been asking and getting support for this purpose for years. […] I asked the community for donations three or four times, for a specific purpose only and with all of the budget spoken for. And after getting the requested amount of money we shut down the donations."

Mirror sites, however, do not need to be noncommercial to enjoy the support of the core LibGen community, they just have to provide free access to users (Bodó 2013; Schultz 2006). This means that ad-supported mirrors are endorsed, but the reselling of texts is frowned upon. The ethical stance of LibGen on this issue is best illustrated via the reconstruction of the conflict with another site, E,[10] which used the LibGen stock to seed its own library and then adopted a "collaborative piracy" business approach.

E is another hugely popular online shadow library, offering access to a million plus titles. It is based on a simple idea: If a user cannot find a book in its collection, the administrators offer to purchase a digital or print copy, rip it, and sell it to the user for a fraction of the original price—typically under $1. Access to E is by invitation only. Payments are made in anonymous Amazon gift cards, which make the purchases easy and protect the identity of the users. E recoups its investment, in principle, through the multiple sales of the same low-priced ripped copy. While clearly illegal, the logic is not that different from that of private subscription libraries, which purchase a resource and distribute the costs and benefits among club members.

Although from the rights holders' perspective there is little difference between the ad-supported and the collaborative piracy approaches, many participants in the pirate librarian community draw a sharp line between the two, viewing the sales model as a violation of community norms. An internal forum post tried to clarify the relationship of LibGen to other services as follows:

E is a scam. They were banned in our forum. Yes, most of the books in E came from LibGen, because LibGen is open, but we have nothing to do with them. […] If you wish to buy a book, do it from legal sources. Otherwise it must be free. […]

Here's what E wants:

- make money on downloads of e-books, no matter what kind.
- get books from all the easy sources, spend as little effort as possible on books, maximize profit.
- no need to build a community, no need to improve quality, no need to correct any errors. Just put all files in a big pile and maximize profit.
- keep files in secret, never give them away, and keep no listing of files so there is no information about what books are really available on E or what is being done.

There are very few similarities in common between E and LibGen, and these similarities are too superficial to serve as a common ground for communication. […]

They [E administrators] run an illegal business, making a profit.

Library Genesis administrators describe a set of values that differentiates possible site models. They prioritize the curatorial mission and the provision of long-term free access to the collection with all the costs such a position implies, such as open sourcing the collection, ignoring takedown requests, keeping a low profile, refraining from commercial activities, and as a result, operating on a minimal budget. E prioritizes the expansion of its catalog on demand, but this implies a commercial operation with a larger budget and the associated higher legal risk. Many of the other sites that mirror LibGen's catalog prioritize public visibility, carry ads to cover costs, but also respond to takedown requests to avoid as much trouble as possible. From the perspective of expanding access, these are not easy or straightforward trade-offs. In LibGen's case, the commitment to the mission of providing free access comes with significant sacrifices, the most important of which is relinquishing control over the shadow library's most valuable asset: its collection of 1.2 million scientific books. But the LibGen admins believe that these costs are justified by the larger goal of making free access independent of the fate of LibGen.

Library Genesis is not the only file sharing community that relies on internal discipline and restraint to ensure the long-term survival of the collection and the community (see, e.g., Bodó 2013). It is unique, however in its radical open source approach. This approach is rooted in the way it regards the legal status of its subject matter—scholarly publications. While openness in the field of scientific research is hardly new, grounded in the understanding that we see further if "standing on the shoulders of giants," LibGen's copynorms are equally shaped by the specificities of the Soviet and post-Soviet era, in which the experiences of repression, scarcity, and expulsion from the first world of scientific knowledge production were paramount.

The Co-development of Copynorms and Copyright Laws in the Post-Soviet Era

The copynorms of the LibGen community were shaped by and reacted to the development of local (Russian) and international laws on the digitization and online distribution of protected works. Russian digital libraries emerged in a period of double transformation: the post-Soviet copyright system had to adopt global norms, while these global norms struggled to adapt to the emergence of digital copying.

The first post-Soviet authors rights law was enacted in 1993. Its major goal was to update the local regulatory framework to conform to at least some international

standards, and to the expectations of Western rights holders, for whom such laws were a precondition for entering the newly opened Russian markets. The first two post-Soviet decades saw significant efforts to harmonize Russian law, at least on paper, with the existing WIPO and World Trade Organization (WTO) frameworks. Yet, significant gaps and uncertainties remained in terms of scope, the legal clarity, or the enforceability of the freshly implemented regulations (Sezneva and Karaganis 2011). This was especially true for rules regarding the digital world. "Internet rights" were introduced only in a 2006 amendment to the authors' rights law (Budylin and Osipova 2007; Elst 2005, 425).

During most of the 1990s, user-driven digitization and archiving took place in a regulatory void where such activities were barely addressed. Under such conditions, informally negotiated norms filled the gap. Authors and publishers who saw their works appear in digital form had to rely on these informal norms to establish control over their texts vis-à-vis enthusiastic collectors and for-profit entrepreneurs. Such regulation via norms did not always work, and it was widely ignored when the subject was foreign work, but for some authors, limited control of a work could be exercised through the copynorms in some of the better-organized Russian file-sharing communities.

The roots of the Russian digital copynorms can be traced back to 1997, when HARRYFAN, an early Russian digital text collection, was first published on CDs. The CD contained around ten thousand texts, consisting mostly of Russian science fiction. It was originally compiled by Igor Zagumenov, a book enthusiast, from works that appeared on the Holy Spirit BBS. The CD was a nonprofit project, which Zagumenov planned to print and sell in a single run of around one thousand copies. Zagumenov contacted some of the authors and publishers, and received permission from some of them to distribute their texts. But the CD also included many other works that were uploaded to the BBS without authorization. In an effort to legitimize the collection, Zagumenov included the following notice alongside his name and contact information and that of the authors who had granted permission: "Texts on this CD are distributed in electronic format with the consent of the copyright holders or their literary agent. The disk is aimed at authors, editors, translators and fans of science fiction and fantasy as a compact reference and information library. Copying or reproduction of this disc is not allowed. For the commercial use of texts please refer directly to the copyright owners at the following addresses."

As the CD circulated, some authors began to notice that their work was used without their authorization. Some complained about the material damage the collection may have caused them, but most focused on moral rather than strictly economic rights: many took issue with the lack of permission, the mutilation of some of the works, the

lack of attribution, and the removal of original copyright and contact notices. Some authors had no problem appearing in the collection per se, but objected to the fact that the CDs were sold (and printed in greater numbers than originally agreed upon in spite of Zagumenov's intentions).

The debate that took place in the book-related fora of Fidonet and drew in a number of the affected authors was useful in revealing and refining community norms. Many participants drew a distinction between the free access provided first by Fidonet and later by lib.ru, which integrated some parts of the Fidonet collection, and what was perceived as Zagumenov's for-profit enterprise—despite the fact that the price of the CD only covered printing costs. The debate also forced authors and publishers to consider the effects of the digital book communities' actions on their business and reputation. Some authors did not want to appear online at all; others wanted only their published works to be circulated, but in any case, the consensus that emerged seemed to agree that online, bottom-up librarianship was beneficial as long as it respected the wishes of the authors.

Moshkov also integrated parts of the HARRYFAN CD into lib.ru. Moshkov's policy toward authors' rights was to ask for permission if the author or publisher could be contacted. He also honored takedown requests. In 1999, he addressed the copyright issues associated with lib.ru:

> The author's interests must be protected on the Internet, including the opportunity to link back to the authorized source, assert the right of attribution, and protect the work from distortion. Anyone who wants to protect his/her rights should be ready to address these problems, ranging from the ability to identify the offending party, to the possibility of proving infringement. […]
>
> Meanwhile, the question how to protect authors-netizens' rights regarding their work published on the Internet has become important. It is known that there are a number of periodicals that reprint material from the Internet without the permission of the author, without payment of a fee, without prior arrangement. Such offenders need to be shamed via public outreach. The "Wall of shame" website is one of the positive examples of effective instruments established by the networked public to protect their rights. It manages to do the job without bringing legal action—relying on polite warnings, indications of potential trouble, and shaming of the infringer.
>
> Do we need any laws for digital libraries? Probably we do, but until then we have to do without. Yes, of course, it would be nice to have their status established as "cultural objects" and have the same rights as a "real library" to collect information, but that might be in the distant future. It would also be nice to have e-library "legal deposits" of publications in electronic form, but when even Leninka [the Russian State Library] cannot always afford that, what we really need are enthusiastic networkers. […]
>
> The policy of Lib.ru is to take everything users give, otherwise they cease to send books. It is also to listen to the authors and strictly comply with their requirements. And it is to grow and prosper. […] I simply want the books to find their readers because I am afraid to live in a world

where no one reads books. This is already the case in America, and it is speeding up with us. I don't just want to derail this process, I would like to turn it around.

Moshkov didn't have answers to all the problems facing authors, but he worked to chart an alternative to both the lack of legal protection and the public cost of a lockdown of digital rights. He played a crucial role in consolidating norms around these practices in Russian digital publishing—a role that was later recognized in various prizes from the International Union for Internet Professionals in Russia.[11] Ultimately, Moshkov's framework rested on the following principles:

- The digitization of books and the practice of online distribution was to be understood to be part of the history and tradition of "the library."
- As is the case with libraries, such practices had to be nonprofit in nature.
- Digital text collections were expected to respect the wishes of the rights holders even if they were not legally obligated to do so.
- Digital librarians were expected to maintain active communication with the different stakeholders in the community, including authors and readers.
- Digital text collections were understood to respond to a clear gap in affordable, legal access.
- Digital texts were not regarded as substitutes for printed books.

Many digital libraries subscribed to Moshkov's principles. But for multiple reasons, by the mid-2000s this consensus was under substantial stress. The latitude that Moshkov had enjoyed was shrinking. Internet and computer access had become mainstream. The legal environment was about to change. But most important, the commercialization of pirate archives had become a viable option and thus a prominent issue for both the community and rights holders.

Formalization of the IP Regime in the 2000s

Russia formally joined the World Trade Organization in 2012. As a condition of membership, Russia had to bring its intellectual property regulation in line with international standards. The road that led to full harmonization started with the first copyright law reform in 1993. Over the next two decades, the United States put unrelenting pressure on the Russian government for further reforms. Throughout the period—and indeed to the present day—U.S. Trade Representative Special 301 reports (which provide a means for U.S. companies to complain about foreign intellectual property [IP] enforcement) described a litany of Russian failures to protect copyright, from inadequate penalties to weak policing to ill-informed judges. Partly in response to these reports, Russia amended

its copyright law in 1998 to the extend the legal framework to encompass digital rights. According to the new rules, digital services had to have a license to distribute digital content on the Internet. The licenses were issued by collecting societies, but the rules did not require that these societies have permission from rights holders, provided that the societies paid royalties to them. The result was a proliferation of collective rights management organizations, competing to license content to digital services (Sezneva and Karaganis 2011). Most of these were regarded as illegal by Western rights holders, who had no contractual relationships with the Russian collecting societies.

The resulting confusion led to many high-profile legal disputes. The best known involved Allofmp3.com, a site that sold music from Western record labels at prices far below those of iTunes or other officially licensed vendors. AllofMP3.com claimed that it was licensed by ROMS, the Russian Society for Multimedia and Internet (Российское общество по мультимедиа и цифровым сетям [НП РОМС]), but ROMS, in turn, was disavowed by Western labels and rights holders. A long legal and diplomatic struggle ensued, leading to a failed criminal prosecution of the site owner and the eventual closure of the site in 2007.

The legal status of online text collections was subject to the same uncertainties and faced similar international pressure. Book piracy was regularly mentioned in Special 301 reports in the 2000s—though the reported losses were small in comparison to the claims of the music, film, and software industries.[12] The regulatory changes implemented in response to the music industry, in any event, affected the digital libraries as well. In most cases, lib.ru relied on direct agreements with authors to make digital texts accessible. However, it also had a license from ROMS to cover works without direct authorization. The outcome of the AllofMP3.com controversy thus had direct consequences on the legality of lib.ru, and for any other digital library that contemplated legalizing its activities through the 1998 licensing scheme.

With a much lower profile and a focus on Russian literary classics, lib.ru avoided the attention of foreign rights holders. It even benefited from state support during the period, receiving a $30,000 grant from the Federal Agency for Press and Mass Communications to digitize the most important works from the 1930s. But the chaotic licensing environment came back to bite Moshkov. In 2005, Moshkov and lib.ru were targeted in a lawsuit brought by an e-book merchant (KM Online), which was trying to establish its own commercial service.[13] The lawsuit was a sign of a slow but significant transformation in the Russian print ecosystem. The first change was economic. The idea of a viable market for electronic books had begun to find a foothold. Electronic versions of texts began to be regarded as potential substitutes for the printed versions, not advertisements for them or supplements to them. Commercial services emerged

that regarded the well-entrenched free digital libraries as competitors, not collaborators. The second change was regulatory. As Russia continued to bring its laws into closer conformance with WTO requirements ahead of its admission, the legal system of protecting authors' rights became more sophisticated and more effective. Russian rights holders could increasingly rely on local laws to enforce their rights. As with KM Online, the same laws enabled many organizations to claim markets and force out competitors—sometimes in ways that amounted to state-backed racketeering (Sezneva and Karaganis 2011). Western rights holders also gained enough power to demand enforcement against RuNet pirate sites. The copynorms negotiated in absence of the law came into conflict with the varying, often contested, and sometimes violent processes of applying the new legal order.

Closure of the Legal Regime

The legal, economic, and cultural conditions under which LibGen and its mirror sites operate today are very different from those of two decades ago. The major legal loopholes are now closed, though according to one shadow librarian, Russian authorities have shown little inclination to pursue LibGen so far:

> I can't say whether it's Russian or Western copyright enforcement that's most dangerous for LibGen; I'd say that Russian enforcement is still likely to tolerate most of the things that Western publishers won't allow. For example, lib.ru and Librusec and other unofficial Russian e-libraries are tolerated even though far from compliant with the law. These kinds of e-libraries could not survive at all in [W]estern countries.

Western publishers have been slow to join record, film, and software companies in their aggressive online enforcement campaigns, and academic publishers even more so. But such efforts are slowly increasing, as the market for digital texts grows and as publishers benefit from the enforcement precedents set or won by the more aggressive rights holder groups. In 2015, LibGen was named as a defendant in an injunction served against pirate book services in a New York court (Bodó 2016). The domain name of one of the LibGen mirror sites, was seized, apparently due to the legal action taken by a U.S. rights holder. Several of the sites now act on DMCA take down notices, removing links to books reported to be infringing (despite the lack of jurisdiction of U.S. law). LibGen has responded to this pressure by receding further into the background, as one anonymous LibGen administrator noted:

> We want books to be available, but only for those who need them. We do not want LibGen to be visible. If one knows where to get books, there are here for him or her. In this way we stay relatively invisible (in search engines, e.g.), but all the relevant communities in the academy know

about us. Actually, if you question people at universities, the percentage of them is quite low. But what's important is that the news about LibGen is spread mostly by face-to-face communication, where most of the unnecessary people do not know about it. (Unnecessary are those who aim [to] profit).

The policy of invisibility is starkly opposed to Moshkov's policy of maximum visibility. LibGen administrators hope that they can survive in the shadows where LibGen can be protected by the Russian academic community:

In Russian academia, LibGen is tacitly or actively supported. There are people that do not want to be included in the archive, but it is hard to say who they are in most cases unless there are DMCA complaints. But in our experience the complainers are only from the non-scientific fellows. […] I haven't seen a single complaint from the authors who should constitute our major problem: professors etc. No, they don't complain. The other complainers are the ever-hungry publishers.

But the protection the academic community has to offer may not be enough to fend off publishers' enforcement actions. LibGen and other shadow libraries responded to the increased legal pressure in a variety of ways (Bodó 2016). They moved the core service further into the darknets. They dropped the domain names under injunction in favor of new ones. They tightened security protocols in their communities. Yet this may not be enough: LibGen and other services face a critical loss of volunteers who are willing to donate time and money and take substantial legal risks to maintain its radically open service. Some of the shadow librarians have already stepped back, having reached the limits of their tolerance for risk. But the larger expectation of the shadow librarians we talked to is that, even if LibGen disintegrates, there will be someone else to carry on: "[I]f people are physically served court orders, they will have to close the site. The idea, however, is that the entire collection has been copied throughout the world many times over. The database is open, the code for the site is open, so other people can continue."

As the other chapters in this volume document, there are innumerable small digitization projects, archiving communities, sharing networks, and distribution channels operating below the enforcement radar, contributing to a constant diffusion of texts and knowledge across geographical, educational, and income boundaries. The Russian shadow libraries are an experiment in whether such efforts can survive at scale. This is clearly no longer a technical question but rather a social and political one, shaped by the balance of forces between publishers, educators, and states. It seems unlikely, at this point, that the big shadow libraries will prompt the creation of new law. Publishers are well behind the other copyright stakeholders in pushing for stronger enforcement, though they are beginning to make more aggressive use of the available tools. By the same token, there is little prospect of a legal accommodation of large-scale

unauthorized distribution of the kind enabled by Library Genesis and its mirrors. But the growth and survival of these sites have a powerful influence on the practices that shape the larger ecosystem, as publishers face pressure on issues of cost and access and as the example of actually existing near-universal libraries pushes academic culture toward open models. The survival of the Russian shadow libraries is an open question, but they can still lose the battle while winning the war.

Notes

1. Michael Hart, the founder of the Gutenberg Project (GP), recalled in his history of the project: "The Bible accounted for all of our successful work in the 1980s except for the preliminary editions of *Alice in Wonderland*. We were working on a *Complete Shakespeare*, but the copyright laws had been changed with so little publicity that we didn't find out about it for years, and thus a huge amount of labor was lost" (Hart 2006).

2. The story of Library Genesis was reconstructed via semistructured interviews with key members of the community, and close reading of the discussions on the closed online forum of the community. Both access to the site and to community members was given under a strict condition of anonymity.

At one point, I shared an early draft of this chapter with interested members and asked for their feedback. Beyond access and feedback, community members helped with the writing of this article by providing translations of some Russian-language source documents, and by reviewing my translations. In return, I provided a small financial contribution to the community, in the value of USD$100.

I reproduced forum entries without any edits to the language, and I edited interviews conducted via instant messaging (IM) services to reflect basic writing standards.

3. See a quantitative analysis in chapter 3.

4. Such sources include collections of fiction, literary works or comics, not collected by LibGen.

5. Such services include automatically providing the same text in different file formats, suited for different e-readers.

6. LibGen is predominantly donation based, while its mirror sites may serve ads or sell documents individually.

7. In comparison, in the United States in 1975 approximately 39,000 new titles were printed (Greco 2005).

8. We share Helen Freshwater's (Freshwater 2003) view that censorship is a more complex phenomenon than the state just blocking the circulation of certain texts. Rather, its modus operandi, institutions, extent, focus, reach, and effectiveness showed extreme variations over time. This short chapter cannot go into this rich history (Alekseeva, Pearce, and Glad 1985; Dewhirst and

Farrell 1973; Ermolaev 1997; Komaromi 2004; Post 1998; Skilling 1989). For our purposes, the key point is that Soviet censorship not only affected literary works, but also extended deeply into scholarly publishing, including natural science disciplines.

9. DJVU is a file format similar to PDF that simplified online book distribution. For books that contain graphs, images, and mathematical formulae, scanning is the only digitization option. However, the large number of resulting image files is difficult to handle. The DJVU file format allows for the images of scanned book pages to be stored in the smallest possible file size, which makes it the perfect medium for the distribution of scanned e-books.

10. Abbreviated to maintain the anonymity of the service.

11. ROTOR, the International Union of Internet Professionals in Russia, voted lib.ru as the "literary site of the year" in 1999, 2001, and 2003; "electronic library of the year" in 2004, 2006, 2008, 2009, and 2010; Moshkov was elected "programmer of the year" in 1999; and "man of the year" in 2004 and 2005.

12. The Special 301 reports cited USD$40 million losses per year to publishers throughout this period, though such estimates were at best a rough guess and by all appearances, a low priority for the USTR. The details, alleged losses, and analysis in these reports changed little from year to year.

13. KMO was an online vendor that sold digital texts for a small fee. Although the KMO collection—like every other collection—had been assembled from a wide range of sources on the Internet, KMO claimed to pay a 20 percent royalty on its income to authors. In 2004, KMO requested that lib.ru take down works by several authors with whom KMO claimed to be in exclusive contract. KMO's claims turned out to be only partly true. KMO had arranged contracts with a number of the heirs to classics of the Soviet period, who hoped to benefit from an obscure provision in the 1993 Russian copyright law that granted copyrights to the heirs of politically persecuted Soviet-era authors. Moshkov, in turn, claimed that he had written or oral agreements with many of the same authors and heirs, in addition to his agreement with ROMS. The lawsuit turned into a major public event, generating thousands of news items both online and in the mainstream press. Authors, members of the publishing industry, legal professionals, librarians, and Internet professionals publicly supported Moshkov, while KMO was generally presented as a rogue operator trying to make easy money on freely available digital resources. Eventually, the court ruled that KMO indeed had one exclusive contract with Eduard Gevorgyan, and that the publication of his texts by Moshkov infringed the moral (but not the economic) rights of the author. Moshkov was ordered to pay 3,000 Rubles (approximately $100) in compensation.

References

Alekseeva, L., C. Pearce, and J. Glad. 1985. *Soviet Dissent: Contemporary Movements for National, Religious, and Human Rights*. Middletown, CT: Wesleyan University Press.

Battles, M. 2004. *Library: An Unquiet History*. New York: W. W. Norton.

Bodó, B. 2013. "Set the Fox to Watch the Geese: Voluntary IP Regimes in Piratical File-sharing Communities." In *Piracy: Leakages from Modernity*, ed. M. Fredriksson and J. Arvanitakis, 241–264. Sacramento, CA: Litwin Books.

Bodó, B. 2016. "Pirates in the Library—An Inquiry into the Guerilla Open Access Movement." Paper prepared for the 8th Annual Workshop of the International Society for the History and Theory of Intellectual Property, CREATe, University of Glasgow, UK, July 6–8, 2016. https://ssrn.com/abstract=2816925 (accessed August 17, 2017).

Borges, J. L. 1998. "The Library of Babel." In *Collected Fictions*, 112–118. New York: Penguin.

Budylin, S., and Y. Osipova. 2007. "Is AllOfMP3 Legal? Non-Contractual Licensing under Russian Copyright Law." *Journal of High Technology Law* 7 (1): 1–18.

Bush, V. 1945. "As We May Think." *Atlantic Monthly* 7:112–124.

Dewhirst, M., and R. Farrell, eds. 1973. *The Soviet Censorship*. Metuchen, NJ: The Scarecrow Press.

Elst, M. 2005. *Copyright, Freedom of Speech, and Cultural Policy in the Russian Federation*. Leiden; Boston: Martinus Nijhoff.

Ermolaev, H. 1997. *Censorship in Soviet Literature: 1917–1991*. Oxford, UK: Rowman & Littlefield.

Freshwater, H. 2003. "Towards a Redefinition of Censorship." *Critical Studies* 22:225–245.

Friedberg, M., M. Watanabe, and N. Nakamoto. 1984. "The Soviet Book Market: Supply and Demand." *Acta Slavica Iaponica* 2:177–192. https://eprints.lib.hokudai.ac.jp/dspace/bitstream/2115/7941/1/KJ00000034083.pdf (accessed August 17, 2017).

Greco, A. N. 2005. *The Book Publishing Industry*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hart, M. S. 2006. "A Brief History of Project Gutenberg." *Michael Hart's Online Writings*. http://hart.pglaf.org/history.06.txt (accessed August 17, 2017).

Komaromi, A. 2004. "The Material Existence of Soviet Samizdat." *Slavic Review* 63 (3): 597–618. doi:10.2307/1520346.

Lessig, L. 2013. "Aaron's Laws—Law and Justice in a Digital Age." Lecture presented at Harvard Law School, Cambridge, MA. http://www.youtube.com/watch?v=9HAw1i4gOU4 (accessed August 17, 2017).

Levin, M. B. 1983. "Soviet International Copyright: Dream or Nightmare." *Journal of the Copyright Society of the U.S.A.* 31:127–162.

Liang, L. 2012. "Shadow Libraries." *e-flux*. http://www.e-flux.com/journal/shadow-libraries/ (accessed August 17, 2017).

Moshkov, M. 1999. "Chto vy vse o kopirayte. Luchshe by *kizhakku* pochitali (Biblioteke *copyright* ne vrag)." *Kompyuter* (300).

Newcity, M. A. 1978. *Copyright Law in the Soviet Union*. New York; London: Praeger.

Newcity, M. A. 1980. "Universal Copyright Convention as an Instrument of Repression: The Soviet Experiment." *Journal of International Law and Economics* 9 (2): 285–324.

Post, R. 1998. *Censorship and Silencing: Practices of Cultural Regulation*. Los Angeles, CA: Getty Research Institute for the History of Art and the Humanities.

Rieusset-Lemarié, I. 1997. "P. Otlet's Mundaneum and the International Perspective in the History of Documentation and Information Science. *Journal of the American Society for Information Science* 48 (4): 301–309.

Ryzhak, N. 2005. "Censorship in the USSR and the Russian State Library." Paper prepared for the IFLA/FAIFE Satellite Meeting: Documenting Censorship—Libraries Linking Past and Present, and Preparing for the Future, The Nobel Institute, Oslo, Norway, August 11–12, 2005. https://www.bibalex.org/wsisalex/8.Censorship%20in%20the%20USSR%20and%20the%20Russian%20State%20Library.doc (accessed August 17, 2017).

Schultz, M. F. 2006. "Copynorms: Copyright and Social Norms." In *Intellectual Property and Information Wealth: Issues and Practices in the Digital Age*, ed. P. K. Yu, 201–236. Westport: Praeger Publishers.

Sezneva, O., and J. Karaganis. 2011. "Russia." In *Media Piracy in Emerging Economies*, ed. J. Karaganis, 149–218. New York: Social Science Research Council.

Skilling, H. G. 1989. *Samizdat and an Independent Society in Central and Eastern Europe*. London: Palgrave Macmillan.

Solzhenitsyn, A. I. 1974. *The Gulag Archipelago 1918–1956: An Experiment in Literary Investigation, Parts I–II*. New York: Harper & Row.

Stelmakh, V. D. 1993. "Reading in Russia: Findings of the Sociology of Reading and Librarianship Section of the Russian State Library." *International Information & Library Review* 25 (4): 273–79.

Stelmakh, V. D. 2001. "Reading in the Context of Censorship in the Soviet Union." *Libraries & Culture* 36 (1): 143–151. doi:10.2307/25548897.