Internet Archive Blogs

A blog from the team at archive.org
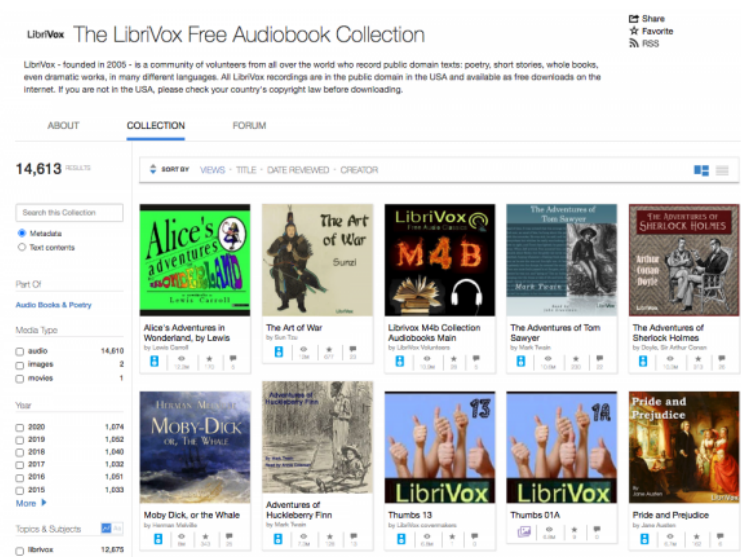
What Information Should we be Preserving in Filecoin?

Posted on October 22, 2020 by Wendy Hanamura

The folks at Protocol Labs love their rockets. And outerspace. And exploration.

So when Filecoin, their cryptocurrency-fueled decentralized storage network launched recently, it was no surprise they called it Filecoin Liftoff. In the payload of that Filecoin rocket are treasures from the Internet Archive:

- 14,613 Free Audiobooks from LibriVox (20 terabytes worth)
- 17,000+ Films from the Prelinger Archive (17 terabytes worth)

For 15 years, LibriVox has harnessed a global army of volunteers, creating 14,200 free public domain audiobook projects in 100 different languages. Where else can you listen to Jules Verne's *20,000 Leagues Under the Sea* in French, Spanish, English, German or Dutch…for free? Now, phrases of Shakespeare, Poe, Joyce and Dante will be stored across the Filecoin mainnet, broken into packets to be reconstituted when needed—perhaps in a new century.

The same destiny awaits the home movies, stock footage, educational and amateur films in the public domain, lovingly curated by the Prelinger Archives founder, Rick Prelinger. He encourages creatives to download and reuse these videos, creating countless new works like this one by musician Jordan Paul:

Now filmmakers and connoisseurs can sleep easier, knowing that a new, distributed copy of those films lives in the Filecoin network, (along with the main copy and multiple backups in the Internet Archive's repositories.)

So what's next Filecoin explorers?

Today, Protocol Labs and the Internet Archive are happy to announce the **Filecoin Archives**, a new community project to curate, disseminate and preserve important open access information often at risk of being lost. You can get involved in so many ways: by nominating information to be stored, uploading it to the Internet Archive, preserving the data as a Filecoin node while earning Filecoin for sharing your storage capacity.

What information should we be preserving? Please tell us!

How about 166,000 public domain books (60 terabytes) from the Library of Congress? Including 2100 texts about Abraham Lincoln and slavery?

Or Open Access Journal articles? (The Internet Archive has collected 9.1 million of them.)

It takes a host of global voices with diverse viewpoints to ensure that humanity's most precious knowledge is represented online and preserved. So we need to hear from you. What open access information or datasets are you interested in preserving?

Between **now and November 5,** please send us your ideas and vote on the others. We will gather your suggestions, add our own, and publish the list from which we will select information to preserve across a global network of Filecoin nodes.

How to send us your suggestions

Look for the tweet from @JuanBenet– reply to it with:

- The Name of the Dataset.
- The size in GB or TB.
- An HTTP or @IPFS link to the data.
- Why it matters.
- #FilecoinArchives

Bonus points if the data is already stored in the Internet Archive or if you upload it there. Vote for ideas by retweeting them and please help us spread the word!

Juan Benet presents his early vision at the 2016 Decentralized Web Summit at the Internet Archive in San Francisco.

In 2015, a young developer named Juan Benet wandered into the Internet Archive headquarters. He painted a picture of a decentralized stack, something he now calls Web3, where the storage, transport and other layers would be distributed across many machines. Together with the DWeb community, we have imagined a web with our values written into the code: values such as privacy, security, reliability, and control over one's own identity. With the launch of Filecoin's mainnet, a piece of that new web is perhaps within reach.

Now it's up to us to make sure the payload includes humanity's most important knowledge.

Posted in Announcements, News | Tagged cryptocurrency, decentralized Web, Filecoin, Librivox, prelinger archives, storage, Web3 | 11 Replies

---

11 thoughts on "What Information Should we be Preserving in Filecoin?"

جوان موزیک
October 22, 2020 at 10:06 pm

i need more information about web3 and filecoin . it is interesting title for me. now i am study about database in TB data.

Pingback: What Information Should We Be Preserving In Filecoin? - RSSFeedsCloud

Gerald William Vaughan
October 26, 2020 at 4:21 pm

Books (print and eBook) published on genealogy, that can be considered references. Hill family/Thomas Jefferson family/George Washington family/Ronald Reagan family.

James Meredith Knox
October 26, 2020 at 4:57 pm

We can still read a 6000 year old scroll or stone tablet. But it's hard to read a 40 year old Scripsit document on a 8″ hard sectored floppy. In addition to documents that SEEM important now, we also need to save engineering information (the tools to use the tools). That means schematics, technical manuals, computer software, etc. Obsolete information that may help also preserve past hardware and information.

Glen Gustafson
October 26, 2020 at 6:27 pm

I deal with the analysis of historical aerial photography, for industrial/environmental cleanup purposes. The idea is, that when a polluted industrial site needs to be cleaned up, somebody has to pay for that extremely expensive process. There may be DOZENS of potential former owners, operators, waste suppliers, insurance companies, etc., disputing about WHEN the damage occurred, and who is responsible. The old, old aerial photography is very helpful in defining many aspects of these issues. To accomplish this, several obscure, uncommon, reference books on the various industries can be helpful. I have several of these. I would like to scan them and make them available to others in the future, after I'm gone. I have been working with aerial photography since beginning as a US Air Force Photo Interpreter in 1962, and founding my consulting company in 1983. These publications are rare and hard to find. Some are declassified Military Intelligence guides; others are training manuals formerly used in this field. The fact that they are now far out of date, is a benefit, because often you find yourself looking at aerial photos from the 1930's through the 1970's. These old reference books are talking about those conditions!

karen frye
October 26, 2020 at 11:15 pm

I would love to see more of the great old radio programs saved with the episodes in order. I loved all those Gunsmoke and Six Shooter episodes but the radio sitcoms are also a hoot. Yes please try to save those, a good alternative to audiobooks.

Sportsclub8
October 29, 2020 at 10:14 am

whats the difference filecoin and bitcoin?

chikavmedia
November 1, 2020 at 4:19 am

The fact is that the Internet will shape the future of many activities and businesses. If we want to put the files in the database of one of the internet libraries as a physical book on the shelf. We need to turn a city into a library. In addition to showing the change in our lifestyle, this shows the great advancement of technology.

David Ballard
November 2, 2020 at 5:56 am

Non-text materials {sheet music, tablatures, symphony scores, choral music) are hard to digitize.

1) list of serial movies (518 or so – how many extant)
2) movie incunabula (video hound golden movie retriever, movie catalogs) and display cards, stills, trailers, variety/trade magazine listings – also "The making of" and "Behind the scenes/On the set'
3) timelines – king lists, synopsis history UK,India,Russia,China, SEA, etc
4) tables like Manetho's lists of pharoahs, Panini's the eight chapters – an explanation of Sanskrit
5) list of DIY/Howto texts broken out like dewey decimal /LC areas of pertinence
6) Interviews – frost/nixon, summit meetings, noted newsreel (Hindenberg, opening concentration camps, World's fair, state fairs, travelogues)
7) Eccentric/fringe like Liberation News Service, small group newsletters
8) catalogs of iconic buildings, scenes, newsfotos – Eiffel Tower, Big Ben, Golden Gate, Pentagon, Blue House, Detroit, Ruhr, aerial views of city streets, parks (Grand Canyon, Igassu Falls, Salto Angel, Rift valley,
hance!

unpin
November 5, 2020 at 7:39 am

Awesome, i like it, and now I'm using archive daily, and it was awesome, also, I'm following archive news few days!

Thanks

Elizabeth Estrin
November 5, 2020 at 8:32 am

The preservation of LGBT materials is imperative for the safety and psychological well-being of LGBT people now and in the future. This material is often not archived elsewhere. I personally have found a dearth of Jewish LGBT materials in particular.

Comments are closed.