

# Internet archivist seeks 1 of every book written

August 1 2011, By MARCUS WOHLSEN , Associated Press

---



In this July 21, 2011 photo, Brewster Kahle, digital librarian for the Internet Archive, demonstrates the zoom function while scanning a book on a touch screen at the Internet Archive's Physical Archive warehouse in Richmond, Calif. Saving a copy of every Web page ever posted sounds like an ambitious life's work, but Kahle has decided digital isn't enough. The founder of the Internet Archive wants to expand his effort to provide "universal access to all knowledge" by preserving a physical copy of every book ever written. (AP Photo/Jeff Chiu)

(AP) -- Tucked away in a small warehouse on a dead-end street, an Internet pioneer is building a bunker to protect an endangered species: the printed word.

Brewster Kahle, 50, founded the nonprofit [Internet Archive](#) in 1996 to save a copy of every Web page ever posted. Now the MIT-trained computer scientist and entrepreneur is expanding his effort to safeguard and share knowledge by trying to preserve a physical copy of every book ever published.

"There is always going to be a role for [books](#)," said Kahle as he perched on the edge of a shipping container soon to be tricked out as a climate-controlled storage unit. Each container can hold about 40,000 volumes, the size of a branch library. "We want to see books live forever."

So far, Kahle has gathered about 500,000 books. He thinks the warehouse itself is large enough to hold about 1 million titles, each one given a barcode that identifies the cardboard box, pallet and shipping container in which it resides.

That's far fewer than the roughly 130 million different books [Google Inc](#) . engineers involved in that company's book scanning project estimate to exist worldwide. But Kahle says the ease with which they've acquired the first half-million donated texts makes him optimistic about reaching what he sees as a realistic goal of 10 million, the equivalent of a major university library.

"The idea is to be able to collect one copy of every book ever published. We're not going to get there, but that's our goal," he said.

Recently, workers in offices above the warehouse floor unpacked boxes of books and entered information on each title into a database. The books ranged from "Moby Dick" and "The Hunchback of Notre-Dame" to "The Complete Basic Book of Home Decorating" and "Costa Rica for Dummies."

At this early stage in the book collection process, specific titles aren't

being sought out so much as large collections. Duplicate copies of books already in the archive are re-donated elsewhere. If someone does need to see an actual physical copy of a book, Kahle said it should take no more than an hour to fetch it from its dark, dry home.

"The dedicated idea is to have the physical safety for these physical materials for the long haul and then have the digital versions accessible to the world," Kahle said.

Along with keeping books cool and dry, which Kahle plans to accomplish using the modified shipping containers, book preservation experts say he'll have to contend with vermin and about a century's worth of books printed on wood pulp paper that decays over time because of its own acidity.

Peter Hanff, acting director of the Bancroft Library, the special collections and rare books library at the University of California, Berkeley, says that just keeping the books on the West Coast will save them from the climate fluctuations that are the norm in other parts of the country.

He praises digitization as a way to make books, manuscripts and other materials more accessible. But he too believes that the digital does not render the physical object obsolete.

People feel an "intimate connection" with artifacts, such as a letter written by Albert Einstein or a papyrus dating back millennia.

"Some people respond to that with just a strong emotional feeling," Hanff said. "You are suddenly connected to something that is really old and takes you back in time."

Since Kahle's undergraduate years in the early 1980s, he has devoted his

intellectual energy to figuring out how to create what he calls a digital version of ancient Egypt's legendary Library of Alexandria. He currently leads an initiative called Open Library, which has scanned an estimated 3 million books now available for free on the Web.

Many of these books for scanning were borrowed from libraries. But Kahle said he began noticing that when the books were returned, the libraries were sometimes getting rid of them to make more room on their shelves. Once a book was digitized, the rationale went, the book itself was no longer needed.

Despite his life's devotion to the promise of digital technology, Kahle found his faith in bits and bytes wasn't strong enough to cast paper and ink aside. Even as an ardent believer in the promise of the Internet to make knowledge more accessible to more people than ever, he feared the rise of an overconfident digital utopianism about electronic books.

And he said he simply had a visceral reaction to the idea of books being thrown away.

"Knowledge lives in lots of different forms over time," Kahle said. "First it was in people's memories, then it was in manuscripts, then printed books, then microfilm, CD-ROMS, now on the digital Internet. Each one of these generations is very important."

Each new format as it emerges tends to be hailed as the end-all way to package information. But Kahle points out that even digital books have a physical home on a hard drive somewhere. He sees saving the physical artifacts of information storage as a way to hedge against the uncertainty of the future. (Alongside the books, Kahle plans to store the Internet Archive's old servers, which were replaced late last year.)

Kahle envisions the book archive less like another Library of Congress

(33 million books, according to the library's website) and more like the Svalbard Global Seed Vault, an underground Arctic cavern built to shelter back-up copies of the world's food-crop seeds. The books are not meant to be loaned out on a regular basis but protected as authoritative reference copies if the digital version somehow disappears into the cloud or a question ever arises about an e-book's faithfulness to the original printed edition.

"The thing that I'm worried about is that people will think this is disrespectful to books. They think we're just burying them all in the basement," Kahle said. But he says it's his commitment to the survival of books that drives this project. "These are the objects that are getting to live another day."

©2011 The Associated Press. All rights reserved. This material may not be published, broadcast, rewritten or redistributed.

Citation: Internet archivist seeks 1 of every book written (2011, August 1) retrieved 10 April 2024 from <https://phys.org/news/2011-08-internet-archivist-written.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.
---