

Community Open Source

# “If you don’t make it beautiful, it’s for sure doomed”: putting the Vault in GitHub’s Arctic Code Vault

GitHub this month installed a massive steel vault, etched with striking AI-generated art, deep within an Arctic mountain, finalizing its Arctic Code Vault. This vault contains the 188 reels of hardened archival film which will preserve the 02/02/2020 snapshot of every active public GitHub repository for 1,000 years. It also now includes a Tech Tree, a human-readable selection of works describing software, computers, and their foundational technologies, along with full-text copies of Wikipedia, Stack Overflow, and other data sources.



Author



Jon Evans

September 20, 2022



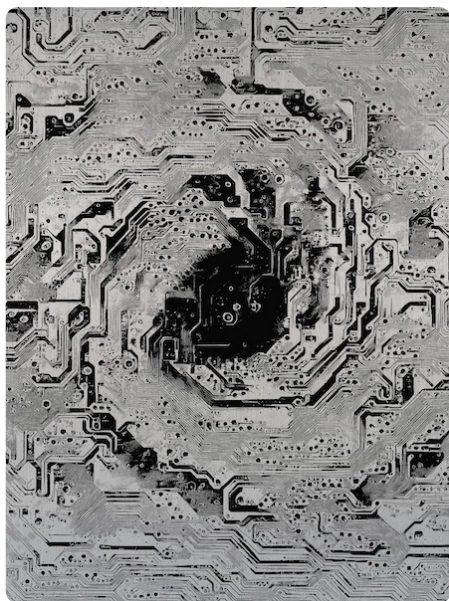
Deep within [an Arctic mountain](#), closer to the North Pole than the Arctic Circle, in a secured and reinforced chamber within a decommissioned coal mine, rest 188 reels of hardened archival film, each

individually sealed against the ravages of time. Encoded on those reels is a snapshot of every active public GitHub repository as of 02/02/2020, the collective work of nearly four million developers. This is the GitHub [Arctic Code Vault](#).

The purpose of GitHub's [Archive Program](#) is to preserve open source software—a hidden cornerstone of modern civilization, and the shared heritage of all humanity—for future generations. Our goal is to secure it for a thousand years. This is an extremely ambitious objective. Few if any of today's tech companies will still recognizably exist a century or two from now.

## A thousand years is a very long time

Ruins that we consider unthinkably ancient, such as Angkor Wat, Great Zimbabwe, and Macchu Picchu, had not yet been built a thousand years ago. The question is not so much whether [the archival reels we deposited two years ago](#) will physically survive, but whether anyone will know or care enough to keep and refer to them. We need to visually signify their importance. That's why [Alexander Rose](#), executive director of our partners, the [Long Now Foundation](#), told us matter-of-factly, "If you don't make it beautiful, it's for sure doomed." That's why the "Greatest Hits" [reels for our institutional partners](#) were enclosed in museum-quality cases, and why this month we installed a massive steel vault, etched with striking AI-generated art created by artist [Alex Maki-Jokela](#), within the Arctic World Archive. GitHub's Arctic Code Vault is now a literal vault, with our archival film reels resting safely within its 1400kg/3000lb edifice. Even if its inheritors many centuries from now don't know what it is, they'll certainly recognize it's something extraordinary.



---

## The Why of The Vault

You may wonder: what do we expect those inheritors to do with it, exactly? There are many plausible

practical purposes. A worrying amount of the world's knowledge is currently stored on ephemeral media: hard drives, CD-ROMs good for a few decades, backup tapes with notional 30-year lifespans.

It's easy to envision scenarios in which an unexpected need arises for software otherwise lost to bit rot, à la the [hunt for mothballed Saturn V blueprints](#) after the Space Shuttle Challenger disaster.

It's also plausible that the value of the archive will be mostly historical. It already acts as something of a pre-pandemic time capsule—the “normal” of 2020 is already greatly changed, only two years later. Future historians may regard our age of open source ubiquity, volunteer communities, and Moore's Law as historically significant. This is one reason why the 02/02/2020 snapshot was so broad and democratic. As another advisor, historian and science-fiction author, [Ada Palmer](#), explained to us, while we have many letters written by the Renaissance's wealthy aristocrats, what modern historians would *really* like—few of which have survived—are ordinary people's shopping lists. Our hope is that by storing and indexing millions of repositories we have captured a valuable cross-section of the world of modern software.



## The Tech Tree

But of course the world does not begin and end with software. Indeed the very concept may be foreign to the archive's inheritors. That is why we have also assembled, archived, and this month deposited into the Code Vault, what we call the "[Tech Tree](#)." While each reel already contains a human-readable [guide](#), in five different languages, describing the archive and its contents, the Tech Tree is a selection of works—mostly human-readable, rather than encoded—which provide much broader context.

The Tech Tree describes how the world makes and uses software today, along with an overview of computers themselves and their foundational technologies. It also includes a small selection of artistic, cultural, and historical works, as cultural context; a snapshot of the full text of [Wikipedia](#), in all of the archive's five languages (Arabic, Chinese, English, Hindi, Spanish); and—of course!—a complete data dump of [Stack Overflow](#) because how could we not?



---

## Onwards and archivewards

With the emplacement of the steel vault in the Arctic World Archive, the deposit of the Tech Tree, our “warm backup” partnerships with the [Internet Archive](#) and [Software Heritage](#), and the “Greatest Hits” reels safeguarded in three world-renowned libraries on three continents, version 1.0 of the GitHub Archive Program has achieved essentially all of its objectives. We don't intend to rest on our archival laurels, though—the Archive Program is an ongoing endeavor, and we look forward to announcing its future initiatives.

---

Tags: [Arctic Code Vault](#), [GitHub Archive Program](#)