

THE ROSETTA DISK

Another tool project by the Long Now Foundation

PAY IT FORWARD

WHILE IT HAS BECOME trivially easy to duplicate information within time—to make millions of copies of things at once—it is still immensely difficult to duplicate information over time—to make sure a copy of something will last centuries or longer. Paradoxically a digital format decreases the longevity of information because digital platforms obsolete so quickly. All the stuff you stored on floppy disks will be unreadable shortly. In a hundred years from now CDs will almost certainly be unreadable because of the presumed widespread lack of operating CD players (think 78 records and eight-track tapes).

Civilization is a tool for remembering, for bringing what was learned in the past into the future. Civilization was hatched by the technology of writing and its attendant technologies of book, index, and library—all meant to keep ideas and information alive over time. Books printed on acid-free paper and stored on shelves do this well. However our expectation of what a book can do has not stood still. The full content of a book—not just its title—must be deeply linked with all other texts and supremely searchable. Any book not on the Web almost doesn't exist now. Thus we are on a campaign to digitize books as fast as we can. But with the proliferation of digital media, and their increasingly short life cycles, maintaining a digital book's knowledge over long periods of time is a new, and vital, problem. The Web itself is no solution. Brewster Kahle, a pioneering archivist at the Internet Archive, has been making the sole backup of the Web. If it goes down (and some publishing and broadcast companies would like it to stop because they consider the backup to be a copyright infringement), then there goes contemporary civilization's only archive.

The Long Now Foundation has been exploring solutions to this problem. Together with the Getty Museum it hosted a "Time and Bits" conference, gathering experts in the conservation of literary and art materials along with computer scientists to ponder how to

conserve digital material over millennia. From that meeting emerged the idea of a 10,000-year library, an institution dedicated to the care and exercise of materials for the very long term. As a first step toward the goal of developing a millennial library, the Long Now Foundation secured funding from the Lazy Eight Foundation to create a prototype of a very long-term storage device. The aim of the prototype is to take a fixed amount of text and embody it in a device that could reasonably be expected to be readable and understood 10,000 years from now.

A number of immediate challenges surface when the problem is stated in this way. What language should the text be in? What format? What material should be used? And what should the content say? The actual goal was a 10,000-volume library, in fully searchable, indexed, multi-linguaged text that would endure a thousand centuries and be readable even if civilization somehow forgot, or gave up, digital technology.

That clearly was too ambitious to start with, so Doug Carlston, co-founder of Broderbund software and Long Now board member, suggested a doable first project as a rehearsal for the grander scheme. His idea: find the most translated texts in the world, gather all the translations and put them all on one super-durable disk. Because all the translations would be parallel the disk could serve as a Rosetta disk. If in the far future a text in a long-forgotten language should be found, there would be at least one source (the 10K Rosetta disk) etched into a noncorroding material that would contain clues to deciphering it. Achieving this modest goal would also teach us a lot about how to do a larger library.

The most translated text in the world at this time is the Bible, in particular the first three chapters of Genesis. The start of Genesis ("In the beginning") is translated into thousands of languages and scripts, and as a bonus, most of these translations were housed in a few institutions. Jim Mason, a young artist with a background in languages and anthropology was hired by Long Now to research and design a 10K Rosetta disk. He imagined a tangible object that would say: read me, keep me forever.

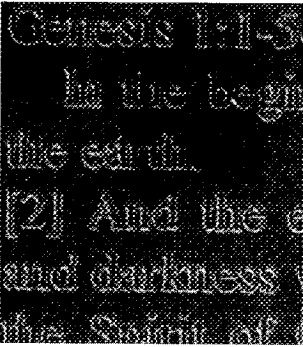
Here is Jim's report on the current status of this project.

We are attempting to create a modern "Rosetta stone" using a new extreme-longevity, high-density analog storage technology. The makers of this storage device, a three-inch micro-etched nickel disk, estimate longevity in the 2,000–10,000-year range. Our goal is to create a functional translation engine for 1,000 world languages as well as an aesthetic object that suggests a journey of the imagination—a tool that both enables and inspires communication between distant historical moments and diverse cultural worlds.

The Rosetta disk expands on the parallel text structure of the original Rosetta stone by archiving five distinct linguistic components for each language on the disk. We have selected these five components as the "minimum representation" necessary for a meaningful recovery of a lost language and/or writing system by future linguistic archaeologists. These components are:

- 1. Main parallel text: We are using the widely translated Genesis Chapters 1-3 for the main text.*
- 2. Indigenous creation myths of each region, including the contemporary creation stories of the Big Bang and human evolution. Each story has an interlinear gloss (grammatical breakdown and analysis).*
- 3. Core vocabulary list.*
- 4. Orthography: The writing system(s) and scripts of the language with pronunciation guide.*
- 5. Metadata for each language: origin and current distribution of language, number of speakers, linguistic family, typology, and history, etc.*

We have selected an extremely stable, nano-scale, analog storage system as an alternative to the quick obsolescence and fast material decay rate of typical digital storage systems. This technology, developed by Los Alamos Laboratories and Norsam Technologies, encodes analog text and images on a 3" nickel disk at densities of up to 350,000 pages per disk. Since the encoding is analog (no 1's or 0's), there is no platform or format dependency, guaranteeing readability despite changes in digital operating systems, applications, algorithms, etc.



The letter "G" is 3 microns across. At this scale one disk can hold 4,000 pages.

Reading the disk requires a microscope, either optical or electron, depending on the density of encoding. The current

version of the Rosetta disk will only require an optical microscope to read. The scan can be combined with a typical Optical Character Recognition system to read the text back into digital formats relevant at the time of reading.



Our current design consists of an Earth image at the center with spokes radiating outward holding 27,000 language data pages, twenty-seven pages for each language. An external band of Genesis texts in eight major world languages (English, Russian, Hindi, Spanish, Hebrew, Mandarin, Arabic, and Swahili) begins at eye-readable scale and slowly tapers down to nano-scale. The design implies the operating instructions: "Get a magnifier because there is more."

The prototype disk is mounted in the center of a 4-inch diameter sphere made of optical glass and stainless steel. Inside the bottom metal hemisphere we have created a hollow cylinder that holds a stainless steel ribbon for disk caretakers to etch their names, locations, and dates—hopefully creating a unique pedigree for each Rosetta object as it travels through time and human hands.

Right now only one copy of the Rosetta disk exists. Ideally tens of thousands of them should be seeded throughout the world as a

means to insure the continuation of their content into the far future. That's one way to keep it alive.

There are four major approaches for keeping things very long term: hide them (Dead Sea Scrolls), make them large and indestructible (pyramids), religiously care for them (relics), or make lots of copies and hope for the best. This last strategy is called LOCKS (Lots of Copies Keep 'em Safe) and it seems the most likely approach for digital information.

But LOCKS doesn't always work. The modern time capsule was invented by Thornwell Jacobs, an educator, minister, and amateur historian in Georgia in the 1930s. Jacobs was disturbed by his studies of ancient biblical history because so much of past records has disappeared. If only people back then had had the wisdom and foresight to keep essential materials for later historians (like him) to study, how much richer our understanding of the world (and religion) would be now. But wait, Jacobs, realized, our time is no better! We also are not considering future historians. We are not deliberately conserving knowledge in a form that would make it easy for them to retrieve and study. If we are so civilized shouldn't we scientifically archive material for the future—and not just the official texts?

Jacobs decided to send to the future as much of civilization as he could squeeze into a large sealed vault. He called it the Crypt of Civilization, a living room-sized underground tomb, and he stuffed it full of encyclopedias, classics, film clips on metallic film, wordless instructions, and a trove of folkloric ephemera and cultural knickknacks. Figuring that there has been about 6,000 years of history, Jacobs set the opening date approximately 6,000 years into the future, or 8113 to be precise according to his ideas of when civilization actually began (4241 B.C.E.). By predefining the opening date, in contrast to treasures hidden in walls and vaults meant to be opened in some accidental time, Jacobs invented a time capsule. (One of his associates took Jacob's idea to the Westinghouse company and got them to build and bury a time capsule several years before Jacobs could finish his own; thus the world remembers the Westinghouse World's Fair one in 1939 as the first time capsule.)

Inspired by the Westinghouse time capsule, every year families, schools, towns, and corporations seal about 100,000 small time capsules in the US. Almost 95 percent of them are lost track of and forgotten in ten years. So how will the world remember the buried Crypt of Civilization in 6,000 years?

Jacobs's solution was a version of LOCKS. He created a book about the Crypt, which was to serve both as an operating manual and as a reminder. The book tells the exact longitude and latitude of the Crypt (under part of Oglethorpe University in the northern outskirts of Atlanta, Georgia), and how to figure out longitude and latitude if both the borders known as Atlanta and the convention of longitude should disappear. Jacobs then made lots of copies of the book, and sent them to libraries around the world, including a few monasteries in Japan and China. The shocking surprise is that after only fifty years or so there are only two known copies of the book.

Forwarding material into the future is difficult. We have lost more texts of the ancient Greeks than we have kept. Manufacturing something to endure thousands of years is only part of the problem. Anticipating changes in language and context is also only part of the problem. Sustaining care is the hard part.

The Rosetta disk, in conjunction with the 10,000-year Clock, is one attempt to develop tools that will provide our technological civilization with a long-term memory. That's a big job that will require many more ideas. For more information, contact www.longnow.org.

—Kevin Kelly