The goal of this project, DNA Storage, is to design and build a user-friendly, cloud-based web application to archive computer files as synthetic DNA strands. The DNA Storage system was built successfully and is accessible at dnastorage.appspot.com. At present, the digital universe is approximately 44 zettabytes (44x1021 bytes) large and the rate of data generation is increasing rapidly. Data storage solutions, however, have remained mostly stagnant. Current disk and solid-state drives are space-inefficient and deteriorate over time; they cannot support this explosion of data. Given that one human’s DNA stores about 100 zettabytes of data and that DNA can preserve information for thousands of years, the ideal solution is to store digital data (represented in binary 0s and 1s) in synthetic DNA strands (oligonucleotides).

To facilitate this process, DNA Storage harnesses the power of this biological phenomenon in a user-friendly web application, creating a superior storage solution. Tests were conducted to evaluate the efficiency of the algorithm, and it was determined that it can store each byte (8 bits) of data in 5.07 nucleotides, which is 100,000 times more dense than current disk-based drives. The DNA Storage website was built using Google App Engine for the user interface, Google Cloud Functions for encoding and decoding, and Google Cloud Storage to give the user access to the encoded and decoded files. The encoding algorithm successfully minimized the frequency of error incurred during synthesis. For proof of concept, encoded DNA files were synthesized in GeneArt High-Q Strings DNA Fragments, and capillary electrophoresis was used to sequence the oligonucleotides in a laboratory.

This complete trial run demonstrates that anyone with browser access can use the DNA Storage system to store files compactly as synthetic DNA strands and accurately retrieve the files after storage.