

WHY THIS MATTERS

Current technologies are energy intensive and may not meet the exponentially increasing demand for data storage. Without alternatives—such as synthetic DNA or glass—businesses, governments, and individuals may lose billions of gigabytes of data in the next decade. However, high costs and slow write speeds of these technologies pose challenges to their wider use.

SCIENCE & TECH SPOTLIGHT:

ALTERNATIVE DATA STORAGE TECHNOLOGIES

Accessible Version

/// THE TECHNOLOGY

What is it? Alternative data storage technologies—such as synthetic DNA and etched glass—are in development to meet growing demand. Current data storage media (e.g., magnetic tapes, DVDs, and hard drives) are likely insufficient to meet the emerging global data storage needs, which are currently estimated to be around 97 trillion gigabytes. Demand is expected to double by 2025.

The plastic and magnetic materials in current storage degrade over time and the technologies become obsolete by newer technology, requiring replacement as often as every 3 years. In addition, researchers estimate that by 2040, 2.4 billion kilograms of wafer-grade silicon—a high-purity component of computer chips and storage devices—would be needed to store the world’s data. Yet the projected supply is estimated to be only 1 percent of demand. Furthermore, current data storage systems require large, energy-intensive facilities to operate and slow degradation of the storage media. Data centers have a significant environmental impact—reportedly consuming about 2 percent of the world’s electricity as of January 2020 and potentially reaching 8 percent by 2030.

How does it work? Synthetic DNA and glass data storage have greater storage capacity and, when stored properly, are more durable than current technologies.

In nature, DNA has been storing information since life began. The same coding system can be used to store digital information in an artificial DNA strand—created in a lab, not by a biological organism. To read the data, established technology known as sequencing can decode DNA (see fig. 1). DNA can hold over 11 trillion gigabytes in a cubic inch of material.

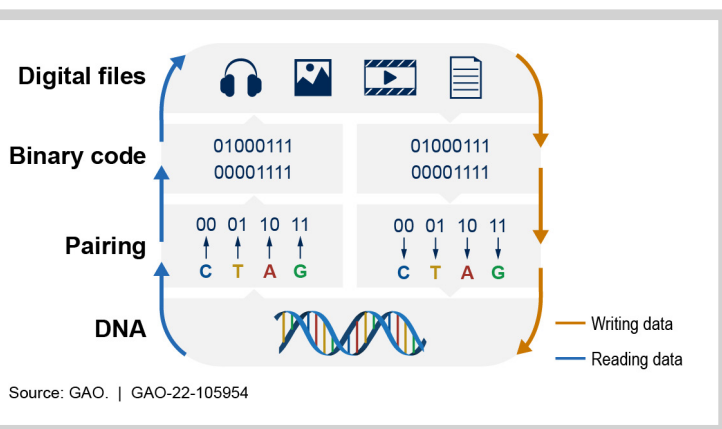


Figure 1. Synthetic DNA data write and read process. The letters A, C, G, and T represent the components of the genetic code.

Data can also be stored in quartz glass using a fast and precise laser, similar to the kind used for vision correction surgery. The laser makes etchings that represent digitally coded ones and zeros. This method is called 5D because it uses five unique attributes of the etchings. Three of the attributes relate to the locations of the etchings on the glass, equivalent to the X, Y, and Z coordinates of a 3D graph. DVD storage uses a similar system, but glass storage has more capacity because the laser creates multiple layers of data using two additional attributes—the size and orientation of the etchings (see fig. 2).

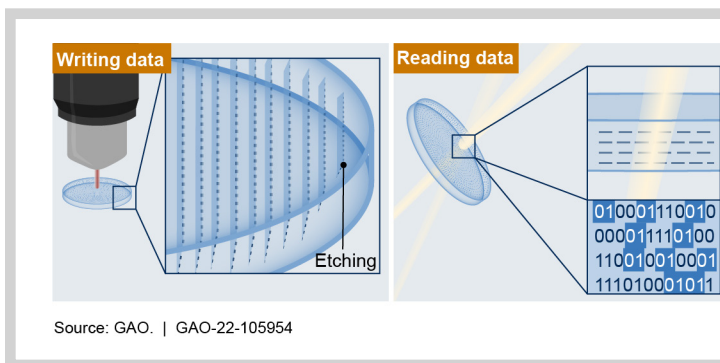


Figure 2. Glass data storage uses a laser to write data and polarized light to read it.

To read the data from the glass, a device shines polarized laser light onto each etching, revealing its five attributes. A camera captures these changes, which a computer decodes into the original digital form. Glass data storage can store hundreds of millions of gigabytes per cubic inch.

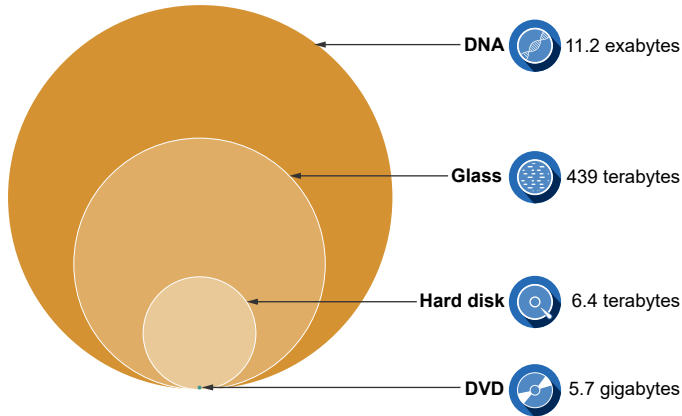
How mature is it? As of April 2022, both of these technologies are still in development, and neither is commercially available. Researchers have successfully used them to store data and some suggest that they could reach commercialization before 2030.

Researchers have successfully stored about 200 megabytes of various forms of data in synthetic DNA, including a music video and the Universal Declaration of Human Rights translated into over 100 languages. Glass storage has been used to store 5 gigabytes of text data in about a 1 inch square of glass. Based on that capacity, researchers believe a CD-sized disk of glass could hold 500 terabytes of data.

Researchers are exploring the capabilities and limitations of each technology. For example, synthetic DNA can store a large amount of data in a very small volume, reducing the size of data storage facilities, but currently costs about \$3,500 per megabyte—millions of times more than silicon-based storage. In addition, this technology’s need for very cold conditions greatly increases its cost. Finally, synthetic DNA data storage is

susceptible to writing and reading errors and researchers are working on correction codes and other methods to reduce errors.

Glass can also store large amounts of data (see fig. 3). However, the speed of writing data is very slow, about 200 kilobytes per second. For comparison, a DVD has a write speed of up to 21 megabytes per second—more than 100 times faster. Glass also presents potential security issues. According to one scientist, some data will remain in the glass unless it is melted or crushed into a powder.



Source: GAO. | GAO-22-105954

Note: Circle area is not linearly proportional to storage capacity; DNA has almost 2 billion times the capacity of a DVD.

Figure 3. Potential data storage capacity by media type (bytes per cubic inch).

Because of these current limitations, both technologies may be best suited for archival data storage rather than everyday applications.

Some federal agencies and companies are interested in the development of these technologies. For example, the Intelligence Advanced Research Projects Activity launched the Molecular Information Storage Program, which aims to use synthetic DNA storage as a low-resource mechanism to meet the data storage needs of the Intelligence Community. Further, at least one large technology company has a project to develop glass data storage as a cost-effective way to meet long-term data storage demands.

/// OPPORTUNITIES

- **Less resource intensive.** Using synthetic DNA and glass for data storage can increase storage capacity, while also reducing the need for some raw materials. It may also result in smaller or consolidated

storage facilities, which could reduce energy use—even with the additional cold storage requirements of synthetic DNA because of its ability to densely store data.

- **Durability.** Data stored in synthetic DNA could last thousands of years if stored at very low temperature. Data stored in glass could last billions of years without degrading, even in harsh conditions.

/// CHALLENGES

- **Enabling technologies.** Synthetic DNA requires specifically designed coding schemes and operating systems to, for example, reduce errors when writing and reading the data.
- **Cost.** Synthetic DNA currently costs millions of times as much as hard drive storage. Some researchers suggest that engineering advances in DNA synthesis are needed to bring down costs. Glass itself is cheap, but components such as lasers and cameras for writing and reading data are expensive upfront costs.
- **Speed.** Glass storage may be limited to archiving data that do not need to be rewritten often due to its very slow writing speeds.

/// POLICY CONTEXT AND QUESTIONS

- What research and development could be pursued to make these alternative data storage technologies more sustainable and affordable?
- What are the design considerations for future data centers that might be able to use these alternative data storage technologies?

/// SELECTED GAO WORK

- Science and Technology: Considerations for Maintaining U.S. Competitiveness in Quantum Computing, Synthetic Biology, and Other Potentially Transformational Research Areas, [GAO-18-656](#).

/// SELECTED REFERENCES

Zhang, J. et al. "5D Data Storage by Ultrafast Laser Writing in Glass" Optical Society of America (2013).

Potomac Institute for Policy Studies. *The Future of DNA Data Storage*. Arlington, VA: 2018.

Church, George M., et al. "Next-Generation Digital Information Storage in DNA." *Science*, vol. 337, no. 6102, 2012, pp. 1628–1628, <https://doi.org/10.1126/science.1226355>.

GAO SUPPORT:

GAO meets congressional information needs in several ways, including by providing oversight, insight, and foresight on science and technology issues. GAO staff are available to brief on completed bodies of work or specific reports and answer follow-up questions. GAO also provides targeted assistance on specific science and technology topics to support congressional oversight activities and provide advice on legislative proposals.

For more information, contact: Brian Bothwell at (202) 512-6888 or BothwellB@gao.gov.

Staff Acknowledgments: Katrina Pekar-Carpenter (Assistant Director), Atiya Siddiqi (Analyst-in-Charge), Jehan Chase, Anika McMillon, Matthew Metz, and Ben Shouse.

This document is not an audit product and is subject to revision based on continued advances in science and technology. It contains information prepared by GAO to provide technical insight to legislative bodies or other external organizations. This document has been reviewed by Timothy M. Persons, PhD, the Chief Scientist of the U.S. Government Accountability Office.

This work of the United States may include copyrighted material, details at www.gao.gov/copyright.