OPPORTUNITY 32

SCOPE VISIONARY

UNCERTAINTIES

Technology, Systems

MEGATRENDS

Boundless Multidimensional Data

TRENDS

Advanced Computing
Bioinformatics
Biotechnology
Data Protection & Privacy
Nanotechnology

SECTORS IMPACTED

Agriculture & Food
Chemicals & Petrochemicals
Communication Technologies & Systems
Consumer Goods, Services & Retail
Cyber & Information Security
Data Science, AI & Machine Learning
Digital Goods & Services
Financial Services & Investment
Government Services
Health & Healthcare
Immersive Technologies
Insurance & Reinsurance
Manufacturing
Materials & Biotechnology
Art, Media & Entertainment
Professional Services
Utilities

What if synthetic DNA met our need for indefinite and unlimited data storage?

DATA HELIX

Synthetic DNA chips meet humanity's increasing need for durable, large-scale data storage, by offering a solution that can preserve information indefinitely while also reducing the environmental footprint of the digital world.

WHY IT MATTERS TODAY

The global datasphere (i.e. the total amount of data in the world) has increased dramatically from 2 zettabytes in 2010 to an estimated 97 zettabytes in 2022[679] and is forecast to increase by 300% by 2025.[680] However, current data storage media degrade over time, leading to data loss or corruption.[681] As data storage capacities continue to grow, managing and organising massive databases becomes complex – effectively retrieving and using such data can prove to be challenging.[682] Unstructured data represent up to 90% of all new enterprise data.[683] These combined factors put the global datasphere at risk of growing decay and inaccessibility. At present, data storage is a massive contributor to climate change: data centres are responsible for 2% of global greenhouse gas emissions.[684]

DNA is being researched as an alternative data storage mechanism. With a storage density of 2.2 petabytes per gram, a DNA hard drive the size of a teaspoon could store all the world's data.[685] These data can be translated into computer-readable files via DNA sequencing. In 2012, geneticists at Harvard University encoded a 52,000-word book in DNA.[686] Researchers at the Eindhoven University of Technology predict the first DNA data centre will be up and running in 5–10 years.[687] Cost decreases will be key to large-scale use – a 2 megabyte file currently costs $7,000 to synthesise and $2,000 to read.[688] Fluorescent labelling of DNA-stored data has been shown to facilitate better data sorting and retrieval.[689]

With a storage density of 2.2 petabytes per gram, a DNA hard drive the size of a teaspoon could store all the world's data

OPPORTUNITY

DNA can keep critical data needed for the functioning of society intact for millions of years[690] in contrast with current data servers, which require constant replacement.[691] Developments in DNA encoding and sequencing can enable speed increases and price decreases, making the technology viable as an everyday enterprise data storage solution. Scientists can develop alternative, less destructive read-write techniques for DNA storage, increasing data durability.

Beyond enhancing data longevity and storage volume, because of their compact footprint and low power demands after data sequencing, DNA data storage centres can have fail-safe measures and multiple backups at minimal economic and environmental cost,[692] reducing data decay and improving sustainability. Taking advantage of DNA's unique structural and biomolecular characteristics can enable next-generation approaches to cryptography and information security.[693]

BENEFITS

In theory, synthetic DNA provides indefinite, durable storage of an unlimited amount of data in less space and without interoperability issues.

RISKS

Accessing data is complex and limited. A lack of procedural and technical guard-rails risks saving massive amounts of erroneous data, leading to inefficiencies in access and use. There are interoperability issues related to historical data and global data sharing with nations where DNA storage is not used. DNA storage continues to be expensive, limiting implementation and potentially exacerbating global inequalities.

The global datasphere has increased dramatically from **2 zettabytes** in 2010 to an estimated **97 zettabytes** in 2022