The Future of Computing Was Inside You All Along: Integrating Your Life's Blueprint with Your Digital Fingerprint

Quilee Simeon

Envisioning the Future of Computing Prize
Social and Ethical Responsibilities of Computing
Massachusetts Institute of Technology

1-Page Summary

The Future of Computing Was Inside You All Along: Integrating Your Life's Blueprint with Your Digital Fingerprint

DNA, often called the "book of life," encodes biological instructions in its sequences of A, T, C, and G nucleotides. Beyond its biological role, DNA's immense data density and stability have positioned it as a revolutionary medium for digital storage. Pioneering efforts by companies like CATALOG have demonstrated encoding entire literary works into synthetic DNA strands, highlighting DNA's potential to surpass traditional storage technologies.

Building on this foundation, the essay envisions a future where digital data is embedded directly into the human genome. Advances in synthetic biology, DNA computing, and gene-editing technologies like CRISPR/Cas9 could enable individuals to securely store personal data—such as medical records or passwords—within their non-coding DNA, offering unmatched privacy, portability, and resilience. Data retrieval would be facilitated through portable sequencing devices, enhanced by encryption for security.

This transformative concept, however, is accompanied by significant challenges. Ethical concerns arise around potential biohacking, unintended mutations, and access inequities. Without proactive measures, DNA-based technologies could exacerbate societal inequalities, granting early adopters disproportionate advantages. Robust regulatory frameworks and equitable access to DNA sequencing tools are essential to mitigate these risks.

Ultimately, merging our biological and digital identities promises to redefine computing by leveraging nature's most efficient storage medium. This integration offers a chance to decentralize data ownership, protect against cyber threats, and align technological innovation with the complexity and dignity of human life.

The Future of Computing Was Inside You All Along: Integrating Your Life's Blueprint with Your Digital Fingerprint

The Book of Life as a Storage Device

When asked what a book is for, most people would say, "to read." Biologists often think of DNA as the "book of life," a collection of strings of characters detailing the instruction manual for building and maintaining organisms. But a book is more than just something to read—it is also a storage device. It contains information encoded in words, just as DNA encodes life's instructions in sequences of nucleotides—the four unique molecular "characters" (A, T, C, and G) that make up DNA. While the scientist-biologist seeks to decode DNA to understand its role as the blueprint of life, the engineer-biologist focuses on leveraging DNA's innate information-storage capabilities for an entirely new purpose: encoding any form of digital information, not just biological instructions.

A company called CATALOG has already demonstrated the commercial feasibility of DNA-based storage by encoding entire literary works into strands of synthetic DNA (see Image 1) [15]. Back in 2022, they wrote the complete works of Shakespeare into less than 1 milliliter of colorless, odorless DNA solution. Today, you can buy Asimov Press' latest anthology in a thumbnail-sized steel capsule of DNA powder [16]. This is possible because DNA is vastly more compact than traditional storage methods. A single gram of DNA can theoretically store up to 455 billion gigabytes of data—equivalent to nearly half a billion 1-terabyte hard drives. Moreover, DNA's molecular nature allows for massively parallel, high-throughput sequencing, enabling rapid and scalable data retrieval. This is computing as nature intended it.

But what if we could take this technology further? Advances in DNA computing show that DNA can store more than just the instructions for life [4]. By embedding our digital data directly into the DNA within our bodies, we could move beyond relying on massive data centers owned by big tech companies. Instead, we would become the sole custodians of our personal information, carrying it within us wherever we go.

This futuristic concept is achievable by integrating gene editing, synthetic biology, and DNA computing. These technologies could enable encoding and retrieving data from the human genome, offering benefits like enhanced privacy, disaster-proof storage, and true data ownership. However, realizing this vision requires addressing ethical concerns such as biohacking, unintended biological consequences, and social inequalities. Ultimately, merging our digital footprint with our biological blueprint offers not just transformative possibilities but a critical chance to shape the future of computing responsibly.

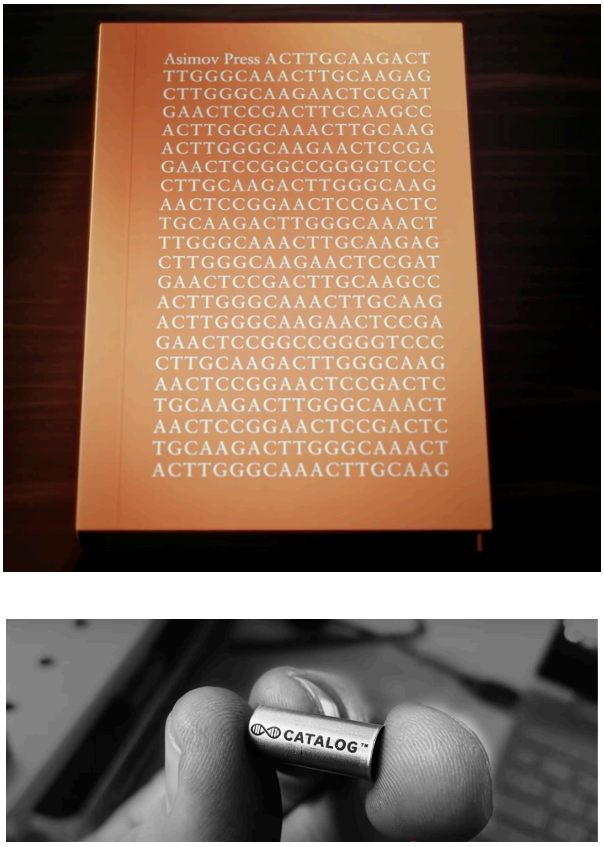

Image 1. Asimov Press transformed its latest book – an anthology of 13 essays spanning hundreds of pages – into DNA. The digital text was encoded into strands of DNA that are millions of base pairs long, using a process developed by CATALOG for translating digital data into the As, Ts, Cs, and Gs of DNA. The book can be bought as a thumbnail sized stainless steel capsule containing the desiccated DNA in white silica powder, making it the first commercially available book written in DNA.

The Science of DNA Computing

Deoxyribonucleic acid (DNA) is a molecule composed of two intertwined strands forming a double helix. These strands consist of four nucleotide bases—adenine (A), cytosine (C), guanine (G), and thymine (T)—which encode the instructions for building and maintaining living organisms. The genome, comprising the complete set of DNA sequences, determines an organism's traits. The Human Genome Project, completed in 2003, mapped the entire human genome, revealing its complexity and potential [8].

Biologically, DNA functions as both a storage system and an executable program. Coding sequences, or genes, are akin to scripts that cells compile into proteins through transcription (DNA to RNA) and translation (RNA to protein). Meanwhile, the vast non-coding regions, once dismissed as "junk DNA," may have essential regulatory or structural roles [9]. These

underutilized regions present an intriguing possibility: a reservoir for data storage that goes beyond DNA's cellular functions.

Synthetic biology, which combines biology, engineering, and computation, has redefined how DNA is viewed. Rather than focusing solely on its biological role, this field treats DNA as a programmable material. Oligonucleotides—short, synthetic DNA sequences—encode digital data by representing binary information in molecular form [17]. These oligonucleotides are assembled into specific sequences using enzymes, creating a molecular "alphabet" that allows vast amounts of information to be stored compactly. This approach bridges biological molecules and digital information systems.

Recent advancements in DNA synthesis and sequencing have made this computational use of DNA possible. Technologies like Illumina's NovaSeq 6000 and Oxford Nanopore's MinION enable faster, more affordable sequencing [10, 11], while rapid cost reductions in DNA synthesis mirror the trajectory of Moore's Law. Though still costly compared to traditional computing, these advancements provide a foundation for scaling DNA storage technologies.

Unlike silicon-based computing, which relies on binary bits, DNA uses a quaternary system with its four bases, achieving unmatched data density [4]. Its durability also makes it ideal for long-term storage, as demonstrated by DNA extracted from ancient specimens [2, 3, 7, 12]. This stability starkly contrasts the short lifespan and obsolescence of modern electronics, which often end up in landfills within a few years [13].

Companies like CATALOG have demonstrated the feasibility of encoding books, films, and datasets into synthetic DNA [15]. Their method involves creating libraries of oligonucleotides that represent digital bits, combining them into sequences, and storing the resulting DNA strands in water or silica powder. When needed, the encoded data can be retrieved through sequencing and translated back into digital form. Advances in portable sequencers, such as the MinION, make this process increasingly accessible [11].

While challenges remain—DNA synthesis and sequencing require specialized tools and expertise—continued innovation could address these barriers. By reimagining DNA as a programmable code, synthetic biology links biology and computing in unprecedented ways. It offers a glimpse into a future where the molecules that define life also serve as the building blocks of our digital world.

Beyond Storage: Merging Digital and Biological Identities

Storing digital data in DNA is already an astonishing concept. But what if we took it a step further? Instead of keeping these DNA-encoded archives in a lab, imagine embedding them directly into the cells of our bodies. Through synthetic biology and advanced gene-editing technologies like CRISPR/Cas9, this futuristic idea could become a tangible reality [14].

CRISPR/Cas9, often referred to as just CRISPR, can be thought of as "molecular scissors" that allow precise edits to DNA sequences. While most current research focuses on targeting the coding regions of the genome to cure diseases or introduce new traits, up to 90% of the human genome consists of non-coding DNA—regions often dismissed as "junk DNA" [9]. Yet, these regions offer an untapped reservoir of genomic "real estate" that could be repurposed to store digital information without interfering with biological functions. With the complete sequencing of the human genome, we already know which regions are safe to modify [1]. CRISPR could be used to embed personalized digital archives in these regions, offering a way to securely store data directly in our cells.

This integration of biological and digital identities opens up a world of possibilities. Imagine a future where your medical records, passwords, or even cherished memories are stored securely within your DNA [18]. Unlike external devices vulnerable to loss, theft, or hacking, your data would move with you, embedded in the one place no external party could access without your consent: your body.

Implementation: How DNA-Based Data Storage Could Work

To understand how this could work, let’s start with a simple example: storing a digital password—perhaps the login credentials for your bank account—within your DNA. The process begins by translating the password into a sequence of nucleotides using a private codebook, a method pioneered by companies like CATALOG. The resulting synthetic DNA strand, known as an oligonucleotide, is barcoded with unique tags to identify it as synthetic and ensure it cannot produce harmful biological proteins.

The next step involves inserting this DNA into a specific, non-coding region of your genome. Using CRISPR/Cas9, a guide RNA sequence directs the Cas9 enzyme to the chosen location, where the DNA is seamlessly integrated. To ensure practicality and safety, the target cells would be somatic—non-heritable cells like those in skin or saliva—so the data stays with you without affecting future generations.

Retrieving the data would be just as precise. For example, extracting a small sample of skin cells would allow scientists to sequence the DNA and isolate the synthetic segment using its barcode. The sequence could then be decoded back into its digital form using the original codebook. Advances in portable DNA sequencers, like Oxford Nanopore’s MinION, already make this process feasible outside of traditional labs, further reducing costs and enhancing accessibility.

Encryption would add an additional layer of security. Just as digital information is encrypted, synthetic DNA could be encoded with barcodes and primer sequences requiring specific guide RNAs to decrypt. This means only someone with the right "key" could access the data, ensuring its privacy even if the DNA were sequenced without authorization.

The potential applications extend beyond passwords. Imagine storing your personal identification, like a DNA-based passport, that could grant access to secure facilities or bank accounts with a quick DNA scan. Unlike conventional forms of identification, this would be nearly impossible to forge, offering unparalleled security in a digital age increasingly plagued by cyberattacks.

Benefits of Storing Data in DNA

DNA offers unparalleled potential as a storage medium due to its extraordinary data density and longevity. Unlike hard drives or silicon-based storage devices that degrade within decades, DNA can remain stable for thousands of years under proper conditions. This durability is demonstrated by the successful sequencing of DNA from woolly mammoths preserved in permafrost for tens of thousands of years, as well as ancient human remains like bog bodies and Neanderthals [2, 3]. Such resilience positions DNA as an ideal candidate for long-term archival storage, offering a stark contrast to the rapid obsolescence of modern electronic devices.

Beyond its longevity, DNA has the potential to revolutionize personal data ownership and privacy. Today, much of our sensitive data—medical records, financial information, and digital identities—are stored in centralized data centers controlled by corporations or governments. Embedding this information directly into an individual's DNA could decentralize data storage entirely. This approach would allow people to carry their most important information within their own bodies, making it immune to hacking, corporate surveillance, or physical theft. For example, instead of relying on external servers, you could access your medical history or banking credentials directly from your DNA using portable sequencing devices.

In addition to privacy, DNA storage offers disaster-proof resilience. Natural disasters, cyberattacks, and geopolitical conflicts threaten traditional data systems, but DNA storage could bypass these risks entirely. Even in extreme scenarios where physical devices fail, DNA-encoded information stored in biological samples could remain intact and retrievable. As portable DNA sequencing technologies like Oxford Nanopore's MinION continue to become more affordable, individuals could theoretically read and access their own data anywhere, reducing dependency on centralized infrastructure.

While these benefits paint an optimistic picture of DNA as the future of data storage, realizing its full potential will depend on overcoming several technical and ethical challenges.

Potential Pitfalls and Ethical Concerns

Embedding digital data into DNA holds extraordinary potential but it is not without risks. Editing the genome raises complex ethical, biological, and social questions. What happens if an unintended mutation occurs? Could bad actors exploit this technology to maliciously alter someone's DNA? These are critical concerns that demand rigorous research, regulation, and oversight before this futuristic idea can be safely implemented. Moreover, the accessibility of DNA-based technology poses significant challenges—without proactive intervention, it risks deepening societal inequalities.

Biological Risks and Unintended Consequences

DNA is often described as a blueprint for life, but our understanding of its complexity remains incomplete. While much of the genome is classified as "junk DNA" because it does not code for proteins, emerging research suggests that some of these regions might have important regulatory or structural roles. Editing these regions, even for non-biological purposes, could produce unforeseen effects that disrupt biological networks in unpredictable ways. For instance, studies in CRISPR gene editing have occasionally resulted in unintended "off-target" edits, raising concerns about unintended consequences when modifying DNA sequences thought to be inert. Such incidents illustrate the delicate balance between scientific potential and biological risk. A well-documented example includes experiments where unintended mutations caused by CRISPR edits led to significant disruptions in cellular behavior.

Biohacking and Security Threats

Storing data in DNA introduces a new realm of security vulnerabilities. As with digital systems, malicious actors could exploit bio-hacking techniques to alter, erase, or even corrupt stored DNA data. For example, synthetic DNA could be targeted by sequences engineered to disrupt stored data, creating a new class of biological malware. Worse yet, such techniques could be weaponized to develop personalized bioweapons targeting specific DNA sequences.

To mitigate these risks, DNA data storage systems must incorporate robust encryption and authentication mechanisms. Similar to how digital information is protected through cryptographic measures, synthetic DNA strands could be barcoded and encrypted with molecular markers that require complementary sequences or proprietary primers to access the stored data. This "nucleic encryption" could ensure that only authorized individuals can retrieve or modify data.

However, such safeguards may not completely eliminate risks, particularly when considering the proliferation of DIY biohacking communities. Today's biohacking often involves at-home CRISPR kits used for genetic experiments, a trend that raises questions about the accessibility and control of DNA editing technologies [5, 6]. Without regulation, this democratization of gene-editing tools could increase the likelihood of misuse.

Economic and Social Inequality

The cost of DNA synthesis, sequencing, and editing has been steadily decreasing, but it remains prohibitive for many individuals. If DNA data storage and editing become commercially available, they could exacerbate existing socioeconomic inequalities. Wealthier individuals would likely gain earlier access to these technologies, enabling them to secure personal data, archive family histories, or even gain genetic enhancements, while marginalized groups are left behind. Such inequities could lead to inherited digital advantages, with families passing down personalized DNA archives across generations, creating knowledge monopolies and further entrenching social divides.

The affordability of DNA readers also plays a critical role. Current portable sequencing devices, such as Oxford Nanopore's MinION, are still expensive for widespread personal use. If these devices remain inaccessible, individuals would be forced to rely on corporations to read and manage their DNA data, raising privacy concerns and creating economic barriers to self-ownership of data. Addressing this issue will require policy interventions that promote the affordability and availability of personal sequencing technologies.

Accessibility and Regulation

For DNA computing to truly revolutionize data storage, it must be accessible to everyone—not just the wealthy or technologically advanced. Achieving this vision requires global cooperation to establish ethical guidelines and regulatory frameworks. These measures must prioritize privacy, safety, and equitable access while minimizing the potential for misuse.

By addressing these concerns head-on, we can ensure that DNA-based data storage advances responsibly, offering its transformative potential to society without creating new vulnerabilities or exacerbating existing inequities.

The Future of Bio-Digital Integration

The integration of DNA computing and synthetic biology represents a significant step forward in technology. For centuries, humans have built machines to assist with tasks, but the potential now exists to embed computational capabilities directly into biological systems. This merging of organic and inorganic components suggests that the future of computing may no longer be limited to silicon chips. Instead, it could be built into the molecular structures that define life, such as DNA.

Synthetic biology, which treats biological systems as programmable materials, is central to this shift. It allows us to reimagine cells, molecules, and organisms as computational units that can store data, perform operations, and even interface with digital systems. This capability opens up opportunities for DNA to not only encode genetic information but also serve as a platform for computation and long-term storage.

Despite the possibilities, this vision raises important questions. How should this technology be governed? Should it be regulated through national policies, international agreements, or decentralized frameworks? How do we ensure that its benefits are shared equitably and not used to exploit individuals? For example, integrating barcoded DNA for identification purposes could improve healthcare and data security but also risks infringing on privacy or reducing people to a data point in a larger system. To move forward responsibly, we need frameworks that protect individual rights and ensure that technological progress aligns with societal values.

Computing as Nature Intended

Advances in DNA computing challenge us to rethink the way we store and process information. For billions of years, DNA has functioned as nature's most efficient and durable storage medium. Today, we are beginning to understand how to adapt this natural system to meet modern needs, such as preserving and protecting digital information.

This technology provides an opportunity to take control of personal data. Instead of entrusting sensitive information to remote servers or external devices, it could be securely stored within our own biological makeup. DNA computing could redefine data security and privacy, giving individuals ownership over their information in a way that was not previously possible.

However, the impact of this technology will depend on the decisions we make now. With careful design and thoughtful regulation, DNA computing could serve as a foundation for secure and accessible data storage. It is not just about advancing technology but about creating systems that respect the complexity and dignity of human life. In doing so, we can build a future where computing integrates seamlessly with biology—leveraging the systems that have supported life for millions of years.

References

1. Lander, E. S., et al. "Initial Sequencing and Analysis of the Human Genome." *Nature*, vol. 409, no. 6822, 2001. DOI: 10.1038/35057062.
2. Slon, V., et al. "Neandertal and Denisovan DNA from Pleistocene Sediments." *Science*, vol. 356, no. 6338, 2017, pp. 605–608. DOI: 10.1126/science.aam9695.
3. Rasmussen, M., et al. "Ancient Human Genome Sequence of an Extinct Palaeo-Eskimo." *Nature*, vol. 463, no. 7282, 2010, pp. 757–762. DOI: 10.1038/nature08835.
4. Erlich, Y., & Zielinski, D. (2017). DNA Fountain enables a robust and efficient storage architecture. Science (New York, N.Y.), 355(6328), 950–954. https://doi.org/10.1126/science.aaj2038
5. Guo, C., Ma, X., Gao, F., & Guo, Y. "Off-Target Effects in CRISPR/Cas9 Gene Editing." *Frontiers in Bioengineering and Biotechnology*, vol. 11, 2023, p. 1143157. DOI: 10.3389/fbioe.2023.1143157.

6. Zettler, P. J., Guerrini, C. J., & Sherkow, J. S. "Regulating Genetic Biohacking." *Science*, vol. 365, no. 6448, 2019, pp. 34–36. DOI: 10.1126/science.aax3248.
7. Green, R. E., et al. "A Complete Neandertal Mitochondrial Genome Sequence Determined by High-Throughput Sequencing." *Cell*, vol. 134, no. 3, 2008, pp. 416–426. DOI: 10.1016/j.cell.2008.06.021.
8. National Human Genome Research Institute. "The Human Genome Project." *Genome.gov*. National Institutes of Health. https://www.genome.gov/human-genome-project.
9. Palazzo, Alexander F., and T. Ryan Gregory. "The Case for Junk DNA." *PLoS Genetics*, vol. 10, no. 5, 2014, e1004351. DOI: 10.1371/journal.pgen.1004351.
10. **Illumina, Inc.** "NovaSeq 6000 System: Powerful Sequencing with Scalable Throughput." *Illumina*. https://www.illumina.com/systems/sequencing-platforms/novaseq.html.
11. **Oxford Nanopore Technologies.** "MinION: Portable Real-Time Device for DNA and RNA Sequencing." *Oxford Nanopore Technologies*. https://nanoporetech.com/products/sequence/minion.
12. Hofreiter, M., et al. "Ancient DNA." *Nature Reviews Genetics*, vol. 2, no. 5, 2001, pp. 353–359. DOI: 10.1038/35072079.
13. Forti, V., et al. "The Global E-waste Monitor 2020: Quantities, Flows, and the Circular Economy Potential." *United Nations University*, 2020.
14. National Institutes of Health. "Gene Editing – Digital Media Kit." *NIH News and Events*. https://www.nih.gov/news-events/gene-editing-digital-press-kit.
15. CATALOG. "Data Economy Meets DNA Computing." *CATALOG*. Accessed February 8, 2025. https://catalogdna.com.
16. **Asimov Press.** "Asimov Press' New Book, Written in DNA." *Asimov Press*, January 2, 2025. https://www.asimov.press/p/technology-book.
17. Park, Hyunjun, and Brian Turczyk. “Synthetic Biology: How to Screen for DNA Danger Now and for the Future.” World Economic Forum, 4 Oct. 2024. https://www.weforum.org/stories/2024/10/synthetic-biology-dna-screening/
18. McFarling, Usha Lee. "Memory Transferred between Snails, Challenging Standard Theory of How the Brain Remembers." Scientific American, 14 May 2018. https://www.scientificamerican.com/article/memory-transferred-between-snails-challenging-standard-theory-of-how-the-brain-remembers/.