



Tecnología de almacenamiento en ADN: una solución sostenible para el Big Data

DNA storage technology: a sustainable solution for Big Data.

Autor/es:

Rommel Antonio Cedeño López ¹

Ing. Rocío Alexandra Mendoza Villamar, Mg. ²

Ing. Ángel Wilson Villarreal Cobeña, Mg. ³



0009-0003-0285-9927



0000-0002-1277-7162



000-0003-0357-0538

¹ SOCKET Studio, Ecuador

² Universidad Laica Eloy Alfaro de Manabí, Ecuador

³ IDrix Technology S.A., Ecuador

rommelcedeno520@gmail.com

rocio.mendoza@uleam.edu.ec

wangelvc@gmail.com

Recepción: 26/03/2025

Revisado: 11/04/2025

Aceptado: 25/04/2025

Publicado: 05/06/2025

Citación/como citar este artículo: Cedeño, R., Mendoza, R. & Villarreal, A. (2025). Tecnología de almacenamiento en ADN: una solución sostenible para el Big Data. V°03 (N°01), Pág. 78-89.

Resumen

La creciente generación de datos a nivel mundial ha planteado importantes desafíos en términos de almacenamiento, sostenibilidad y eficiencia energética. En este contexto, la tecnología de almacenamiento en ADN surge como una solución innovadora y sostenible para enfrentar el problema del Big Data. La justificación de esta investigación radica en la necesidad de buscar medios alternativos y duraderos que superen las limitaciones de los sistemas de almacenamiento tradicionales, como los discos duros y servidores, que consumen grandes cantidades de energía y requieren constante mantenimiento. El objetivo de este estudio es analizar el potencial del ADN como medio de almacenamiento masivo, evaluando su capacidad, durabilidad, densidad de información y viabilidad técnica. Los materiales y métodos empleados incluyen una revisión bibliográfica de investigaciones recientes sobre codificación de datos digitales en secuencias de bases nitrogenadas (A, T, C, G), síntesis y secuenciación de ADN, así como análisis comparativo con tecnologías convencionales de almacenamiento. Los resultados evidencian que el ADN puede almacenar hasta 215 petabytes por gramo, con una estabilidad de miles de años en condiciones adecuadas. Aunque los costos actuales son elevados, se ha observado un progreso acelerado en la reducción de precios y tiempos de lectura/escritura.

Palabras claves: ADN, almacenamiento de datos, Big Data, sostenibilidad, biotecnología, tecnología emergente.

Abstract

The growing global generation of data has posed significant challenges in terms of storage, sustainability, and energy efficiency. In this context, DNA storage technology emerges as an innovative and sustainable solution to address the Big Data problem. The justification for this research lies in the need to find alternative and durable media that overcome the limitations of traditional storage systems, such as hard drives and servers, which consume large amounts of energy and require constant maintenance. The objective of this study is to analyze the potential of DNA as a mass storage medium, evaluating its capacity, durability, information density, and technical feasibility. The materials and methods employed include a literature review of recent research on digital data encoding in nitrogenous base sequences (A, T, C, G), DNA synthesis and sequencing, as well as a comparative analysis with conventional storage technologies. The results show that DNA can store up to 215 petabytes per gram, with a stability of thousands of years under appropriate conditions. Although current costs are high, rapid progress has been observed in reducing prices and read/write times.

Keywords: DNA, data storage, Big Data, sustainability, biotechnology, emerging technology.

Introducción

En la era digital contemporánea, la producción de datos crece a un ritmo exponencial. Se estima que para el año 2025 se generarán más de 175 zettabytes de información a nivel mundial, lo que plantea desafíos sin precedentes en cuanto a almacenamiento, acceso y sostenibilidad de los datos. Las tecnologías actuales, como los discos duros magnéticos y las unidades de estado sólido, si bien han avanzado considerablemente, presentan limitaciones en cuanto a capacidad, durabilidad y consumo energético. En este contexto, surge un enfoque revolucionario que se inspira en la biología: el almacenamiento de datos en ácido desoxirribonucleico (ADN).

El ADN, más allá de ser el portador de la información genética de todos los seres vivos, posee una densidad de almacenamiento sin igual: un solo gramo puede contener hasta 215 petabytes de datos. Además, su estabilidad a largo plazo y resistencia a condiciones ambientales extremas lo convierten en una solución prometedora para el almacenamiento masivo de información digital. Diversas investigaciones han demostrado que es posible codificar archivos digitales, como textos, imágenes, audios e incluso videos, en secuencias de ADN sintético utilizando algoritmos específicos de codificación y decodificación.

El propósito de esta revisión es analizar el estado actual de la tecnología de almacenamiento en ADN, sus fundamentos científicos, los avances más relevantes en el área, sus ventajas comparativas frente a los medios tradicionales y los desafíos técnicos y éticos que aún deben resolverse. Asimismo, se evaluará el potencial del ADN como solución sostenible para enfrentar las demandas del Big Data, considerando tanto sus implicaciones tecnológicas como su impacto ambiental y económico. Esta revisión pretende servir como una guía comprensiva para investigadores, ingenieros, académicos y responsables de políticas tecnológicas interesados en el futuro del almacenamiento de información.

Desarrollo

El desarrollo de la tecnología de almacenamiento en ADN ha sido impulsado por investigaciones interdisciplinarias que combinan biotecnología, informática, química sintética y ciencia de materiales. A continuación, se presentan y analizan algunos de

los estudios más representativos en el campo, sus principales hallazgos, así como las diferentes perspectivas teóricas que los sustentan.

Estudios Clave y Avances Recientes

Uno de los primeros hitos fue el trabajo de (Church, G. et al, 2012), quienes demostraron por primera vez que era posible codificar texto digital en secuencias de ADN sintético. Utilizando un sistema binario para convertir los datos en nucleótidos (A, T, C, G), lograron almacenar una copia del libro *Regenesis* en una cadena de ADN. Este estudio abrió la puerta a la idea de que el ADN no solo es un portador biológico de información, sino también un medio físico para almacenar datos digitales.

En 2017, investigadores de Microsoft Research y la Universidad de Washington lograron almacenar 200 megabytes de datos —incluyendo imágenes, videos y documentos— en cadenas de ADN. Utilizaron un sistema automatizado de codificación y síntesis de oligonucleótidos, mejorando significativamente la eficiencia y reduciendo el margen de error en la lectura y escritura de datos. Este proyecto sentó las bases para el desarrollo de un sistema escalable de almacenamiento molecular.

Otro estudio notable es el de (Abram, K. Z., & Udaondo, Z. et al, 2023), quienes propusieron el algoritmo DNA Fountain, inspirado en las fuentes de datos de redes. Este método mejoró la eficiencia de codificación al reducir la redundancia y los errores durante la síntesis y la lectura del ADN. Lograron almacenar una película completa (*Arrival*, 2016), un sistema operativo y una tarjeta de regalo de Amazon, mostrando la versatilidad del ADN como soporte de información.

Más recientemente, en 2020, un equipo del European Bioinformatics Institute (EBI) logró codificar datos en ADN con una eficiencia cercana al límite teórico de almacenamiento, utilizando herramientas de compresión y técnicas de secuenciación de última generación como Nanopore y Illumina. Además, se exploraron métodos de encapsulación del ADN en esferas de sílice para preservar su integridad durante miles de años.

Implicaciones y Retos

Los estudios analizados coinciden en que el ADN ofrece ventajas inigualables frente a los métodos tradicionales:

1. Alta densidad: 1 gramo de ADN puede almacenar hasta 215 petabytes.
2. Estabilidad: puede durar más de mil años sin degradación si se conserva adecuadamente.
3. Sostenibilidad: requiere menor consumo energético comparado con los centros de datos actuales.
4. Sin embargo, también se identifican desafíos técnicos importantes:
5. Costos de síntesis y secuenciación: aunque están disminuyendo, siguen siendo prohibitivos para aplicaciones masivas.
6. Velocidad de lectura y escritura: la codificación en ADN es mucho más lenta que en los discos duros convencionales.
7. Errores en el proceso: la síntesis química y la secuenciación pueden introducir errores que deben corregirse con algoritmos robustos de verificación y redundancia.

Perspectivas Teóricas

Desde una perspectiva computacional, el ADN se puede considerar como un sistema de almacenamiento no volátil de ultra alta densidad. Los modelos teóricos proponen que puede superar los límites físicos del silicio y otros materiales tradicionales. En este sentido, el modelo de Shannon para la codificación de información se ha adaptado para trabajar con secuencias genéticas, optimizando la redundancia y la entropía del mensaje.

Desde el punto de vista de la biotecnología, la síntesis de ADN artificial ha avanzado gracias a la química de fosoramidita y los procesos de secuenciación de tercera generación. Además, teorías de encapsulación molecular y nanotecnología se han sumado al diseño de sistemas de protección para el ADN sintético, como microcápsulas que lo aíslan de la humedad, luz y oxígeno.

Finalmente, en la perspectiva socioambiental, diversos autores como (Nguyen et al., 2020), destacan el potencial ecológico de esta tecnología. A diferencia de los centros

de datos tradicionales, que requieren refrigeración constante y grandes cantidades de energía, el almacenamiento en ADN podría funcionar en condiciones ambientales normales y sin mantenimiento por largos periodos.

Codificación y Decodificación de Datos en ADN: Principios y Avances

Una de las fases más cruciales en el almacenamiento en ADN es el proceso de codificación de los datos digitales en secuencias de nucleótidos. Cada archivo digital debe convertirse en un formato binario, luego traducirse a un alfabeto de ADN (A, T, C, G) mediante algoritmos que eviten repeticiones problemáticas y aseguren una lectura precisa. Uno de los algoritmos más avanzados es el DNA Fountain, que permite una codificación cercana al límite de eficiencia de Shannon, con una tasa de éxito de recuperación de datos del 100%.

Otro enfoque importante es el de (Adleman, 1994), quienes desarrollaron un sistema de corrección de errores basado en códigos Reed-Solomon. Esto les permitió almacenar y recuperar datos con una durabilidad proyectada de más de 2,000 años, al encapsular el ADN en vidrio y conservarlo en condiciones estables. Esta propuesta demostró que no solo es posible almacenar datos, sino también garantizar su fidelidad a largo plazo, lo que representa un avance fundamental para archivos históricos y científicos.

Por otro lado, estudios como el de (Organick, 2018) demostraron que es posible realizar almacenamiento aleatorio (random access) en ADN, es decir, acceder a fragmentos específicos de información sin necesidad de leer todo el archivo. Esto fue logrado mediante el uso de cebadores (primers) únicos que dirigen la amplificación PCR hacia secuencias particulares, imitando la estructura de carpetas o bloques de datos de los sistemas modernos.

Estos avances apuntan a una mejora sustancial en la velocidad de producción de ADN sintético y abren la posibilidad de fabricar unidades modulares para centros de datos moleculares.

Nuevas Tendencias: Almacenamiento In Vivo y Computación Molecular

Más allá del ADN sintético, algunas investigaciones han comenzado a explorar el almacenamiento in vivo, es decir, dentro de organismos vivos. Por ejemplo, (Shipman, 2019), lograron insertar una imagen codificada digitalmente en el genoma de *Escherichia coli*, que luego fue recuperada mediante secuenciación. Este experimento fue una prueba de concepto que combinó ingeniería genética con tecnologías de información.

Además, la computación molecular propone que el ADN no solo sirva para almacenar, sino también para procesar información. En este sentido, estudios de Adleman y otros en el campo del DNA Computing han demostrado que las cadenas de ADN pueden usarse para resolver problemas matemáticos complejos, como el problema del viajante (TSP), mediante reacciones bioquímicas.

Sostenibilidad y Comparación con Tecnologías Actuales

Uno de los puntos más atractivos del almacenamiento en ADN es su potencial para reducir el impacto ambiental. Mientras que los centros de datos tradicionales requieren refrigeración constante, infraestructuras de alta demanda energética y ciclos de actualización tecnológica cada pocos años, el ADN puede preservarse durante milenios con un mantenimiento mínimo.

Desafíos Éticos, Legales y de Bioseguridad

La posibilidad de usar ADN como medio de almacenamiento ha despertado también debates éticos y legales. Por ejemplo, ¿qué ocurre si se utilizan secuencias genéticas reales? ¿Existe el riesgo de que estas contengan información biológica sensible? Algunos expertos, como (Church, 2019), han advertido que la regulación del almacenamiento genético debe evitar confusiones entre datos sintéticos y biológicos para prevenir riesgos de bioseguridad o mal uso.

A su vez, la privacidad y propiedad de los datos almacenados en material biológico genera un nuevo campo de discusión legal. La posibilidad de que el almacenamiento en ADN se utilice en aplicaciones militares, de vigilancia o de espionaje también plantea cuestiones relacionadas con los derechos digitales y humanos.

Perspectivas Futuras

A corto plazo, se espera que el almacenamiento en ADN se aplique en áreas específicas que requieren conservación de datos a muy largo plazo, como:

- Archivos históricos y culturales.
- Conservación de conocimiento científico.
- Almacenamiento en condiciones extremas (espacio, desiertos, subsuelo).
- Respaldo de registros médicos o legales en instituciones gubernamentales.

A mediano y largo plazo, con la reducción de costos y el aumento en la velocidad de escritura, el ADN podría convertirse en un componente complementario en la infraestructura digital mundial, trabajando junto a tecnologías clásicas como la computación cuántica o la nube híbrida.

Análisis de resultados

La revisión bibliográfica realizada evidenció que la tecnología de almacenamiento en ADN representa un avance significativo en la gestión de grandes volúmenes de datos. Los principales resultados encontrados en los estudios analizados fueron los siguientes:

1. Alta densidad de almacenamiento: Se determinó que un solo gramo de ADN es capaz de almacenar hasta 215 petabytes de datos digitales, superando ampliamente las capacidades de los sistemas de almacenamiento convencionales.
2. Durabilidad excepcional: El ADN puede conservar la información durante miles de años si se mantiene en condiciones ambientales adecuadas, como baja humedad y temperaturas controladas. Esto fue demostrado mediante técnicas de encapsulación en materiales como sílice o vidrio.
3. Avances en codificación y decodificación: Se identificaron algoritmos altamente eficientes, como DNA Fountain, que permiten una codificación optimizada de datos digitales en secuencias de nucleótidos, reduciendo errores y mejorando la recuperación de la información.

4. Eficiencia energética: A diferencia de los centros de datos tradicionales, el almacenamiento en ADN no requiere energía constante para su mantenimiento, lo que lo convierte en una solución más sostenible y ecológica.
5. Viabilidad técnica demostrada: Investigaciones recientes han logrado almacenar diversos tipos de archivos (textos, imágenes, videos, sistemas operativos) en cadenas de ADN, validando su funcionalidad mediante procesos de síntesis y secuenciación.
6. Avances en acceso a datos: Se ha demostrado la posibilidad de realizar acceso aleatorio a fragmentos específicos de información almacenada en ADN, utilizando primers específicos en técnicas de amplificación por PCR.
7. Progreso en reducción de costos: Aunque actualmente los costos de síntesis y lectura de ADN son elevados, la tendencia muestra una disminución progresiva gracias a la automatización y mejoras en biotecnología.

Los resultados obtenidos a partir del análisis de la literatura científica evidencian que la tecnología de almacenamiento en ADN constituye una alternativa altamente prometedora frente a los desafíos actuales del Big Data. Su capacidad de almacenamiento, durabilidad y eficiencia energética la posicionan como una solución disruptiva ante las limitaciones de los sistemas tradicionales, como discos duros, unidades de estado sólido y servidores.

La alta densidad informativa del ADN, con la posibilidad de almacenar hasta 215 petabytes por gramo, permite concebir una nueva generación de soportes digitales ultra compactos. Esta característica es especialmente relevante en un contexto global donde la producción de datos crece exponencialmente. Asimismo, la estabilidad a largo plazo que ofrece el ADN, al mantenerse funcional por miles de años bajo condiciones adecuadas, representa una ventaja crucial para la conservación de información histórica, científica y legal.

El análisis también revela avances notables en los algoritmos de codificación y decodificación de datos, como el DNA Fountain, que optimizan la fidelidad y eficiencia del proceso de almacenamiento. No obstante, persisten desafíos técnicos, especialmente en cuanto a la velocidad de lectura/escritura y los costos asociados a la síntesis y secuenciación, que aún limitan su aplicación a gran escala. A pesar de ello, se ha identificado una tendencia clara hacia la reducción de estos costos,

impulsada por el desarrollo de tecnologías automatizadas y técnicas de secuenciación de última generación.

En términos de sostenibilidad, el almacenamiento en ADN representa una opción considerablemente más ecológica que las infraestructuras de centros de datos convencionales, que requieren un consumo energético constante para refrigeración y funcionamiento. Esta ventaja cobra especial relevancia en el contexto de la crisis ambiental y la necesidad de adoptar soluciones tecnológicas sostenibles.

Finalmente, la revisión también permite reflexionar sobre los aspectos éticos y legales que surgen del uso del ADN como soporte informático. La posibilidad de utilizar secuencias genéticas reales o insertarlas en organismos vivos plantea interrogantes sobre la bioseguridad, la privacidad de la información y la regulación internacional que debe acompañar el desarrollo de esta tecnología emergente.

En conjunto, los hallazgos revisados muestran que, aunque aún en etapa de desarrollo, el almacenamiento en ADN presenta el potencial para transformar radicalmente la forma en que se gestionan y conservan los datos a largo plazo, especialmente en ámbitos críticos donde la integridad, longevidad y sostenibilidad de la información son esenciales.

Conclusiones

La presente revisión permitió identificar que el almacenamiento de datos en ADN representa una solución innovadora y sostenible frente a los crecientes desafíos del Big Data. Su capacidad de almacenamiento masivo, durabilidad milenaria y bajo consumo energético lo posicionan como una alternativa viable para conservar información digital de forma segura y eficiente a largo plazo. Aunque los costos actuales y la velocidad de escritura y lectura continúan siendo limitaciones importantes, las investigaciones revisadas evidencian un progreso sostenido en la superación de estos obstáculos mediante avances en biotecnología, algoritmos de codificación y automatización de procesos.

El ADN ha demostrado ser un medio versátil y confiable no solo para preservar textos e imágenes, sino también para almacenar sistemas operativos, videos y otros tipos de archivos complejos. Asimismo, las posibilidades de acceso aleatorio y de

encapsulación segura amplían su aplicabilidad a entornos extremos o de difícil acceso, como el espacio exterior o archivos históricos de larga conservación. No obstante, la implementación masiva de esta tecnología requerirá el desarrollo de marcos regulatorios que consideren los riesgos éticos y de bioseguridad asociados al uso de material genético sintético o real.

Se recomienda fomentar la investigación multidisciplinaria que integre informática, biotecnología, nanotecnología y derecho para impulsar la evolución de esta tecnología emergente. A futuro, el ADN podría integrarse como parte complementaria de las infraestructuras digitales globales, aportando no solo capacidad técnica sino también sostenibilidad ambiental en la gestión de datos a escala planetaria.

Referencias

- Abram, K. Z., & Udaondo, Z. et al. (18 de Septiembre de 2023). Biotecnología microbiana. Obtenido de Biotecnología microbiana: <https://doi.org/10.1111/1751-7915.14291>
- Adleman, L. M. (1994). Cálculo molecular de soluciones a problemas combinatorios. *Ciencia. Scientia*, 1021–1024.
- Church, G. et al. (2012). *Ciencia Tecnológica*. Obtenido de *Ciencia Tecnológica*: <https://doi.org/10.1126/science.1226355>
- Church, G. M. (2019). Almacenamiento de información digital de próxima generación en el ADN. *Ciencia de la Biotecnología*.
- Nguyen et al. (2020). Arquitectura de centros de datos para la sostenibilidad: Almacenamiento de datos más ecológico con ADN sintético. . *La electrónica se vuelve ecológica en 2020*.
- Organick, L. e. (2018). Acceso aleatorio en el almacenamiento de datos de ADN a gran escala. . *Naturaleza Biotecnología*, 242–248.
- Shipman, S. L. (2019). Codificación CRISPR-Cas de una película digital en el genoma de una población de bacterias vivas. . *Ciencia Biotecnológica*.
- Song, X., & Reif, J. (29 de Julio de 2019). Desafíos y perspectivas. *ACS Nano*. Obtenido de *Desafíos y perspectivas. ACS Nano*: <https://pubs.acs.org/doi/10.1021/acsnano.9b02562>