

DNA data storage

"Store Digital Data in DNA"

Cheryl KOUADIO

National School of Statistics and Data Analysis

September 16, 2023



Contents

- 1 A brief history of digital storage
- 2 Why do we want to store data in DNA
- 3 How can we store digital data to DNA ?
- 4 Limits
- 5 Litterature review

A brief history of digital storage(1/2)

- Since 1947, people have been finding ways to store : from the Williams-Kilburn tube to CFexpress Cards in 2016 (through hard drive, cloud storage etc.)
- Approximatively 23 ways found
- We are finding for new ways to store data
- But why ?

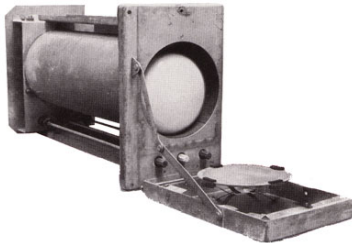
A brief history of digital storage(1/2)

- Since 1947, people have been finding ways to store : from the Williams-Kilburn tube to CFexpress Cards in 2016 (through hard drive, cloud storage etc.)
- Approximatively 23 ways found
- We are finding for new ways to store data
- But why ?

These plenty ways to store digital no longer meet our storage need

A brief history of digital storage(2/2)

Goal throughout all these years : How to store much data in less material ?



(a) The Williams Tube(1947)
Storage capacity : 512 to 1024 data bits

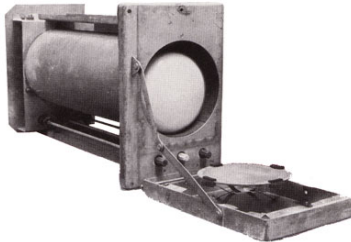


(b) CFexpress card(2016)
Storage capacity : 256 GB to 512 GB data

A brief history of digital storage(2/2)

Goal throughout all these years : How to store much data in less material ?

Solution found : store digital data in DNA



(a) The Williams Tube(1947)
Storage capacity : 512 to 1024 data bits



(b) CFexpress card(2016)
Storage capacity : 256 GB to 512 GB data

Why do we want to store data in DNA(1/2)

- DNA, or deoxyribonucleic acid, is a molecule that carries all informations about human being without electricity.
- In a human, the total weight of DNA = 6 grams.
- Its total storage capacity is on the order of zettabytes (1 zettabyte = 1 trillion gigabytes) <—**HUGE !!!** .
- An example of an immense storage capacity in a molecule .

Why do we want to store data in DNA(2/2)

Think of it as a very special kind of **data structure** :

- Data Containers (super advanced and natural)
- Letters and Instructions ('A', 'T', 'C', 'G' in DNA is a bit like using '1s' and '0s' in a computer)
- Reading and Using Data
- Organization (for better chance of recovering)
- Long term storage (can last for thousands of years)

How can we store digital data to DNA ?

Three main steps :

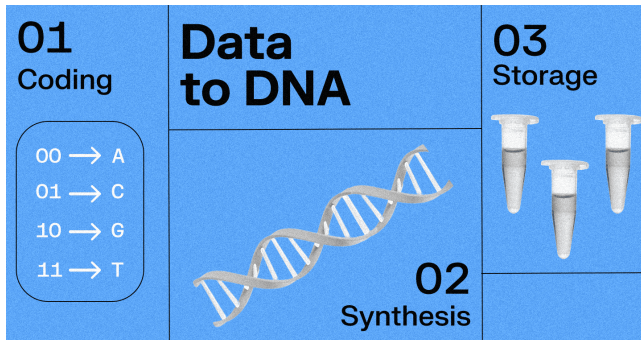


Figure: Illustration from western digital.com

<https://www.youtube.com/watch?v=LTnJWxC03M4>

Focus on step 2 & 3 : synthesis and storage

Through coding, we obtain a special sequence of nucleic bases.

- Synthetic DNA is built one base at a time using chemical synthesis methods.
- Special chemicals and chemical reactions are used to add each base, step by step, to construct the DNA **molecule**.
- Quality control checks are conducted to verify that the DNA sequence matches the intended design

Recovering data through sequencing

- To retrieve the data, the segments are **sequenced** (“read”) and then **decoded** back into the original file (methods of data collection and data analysis).
- One feature of the indexing system is it allows part of the file to be biologically recovered (“random access”) before sequencing, so only data of interest is sequenced.
- All data is recovered error-free because error-correcting algorithms are used during the encode/decode process.

Limits of this technology

- High cost
- Time consuming
- Environmental Sensitivity
- Biological Contamination

Litterature review

- https://en.wikipedia.org/wiki/DNA_digital_data_storage
- <https://www.computerhistory.org/timeline/memory-storage/>
- <https://www.cnrs.fr/en/how-can-digital-data-stored-dna-be-manipulated>
- https://en.wikipedia.org/wiki/DNA_synthesis
- <https://blog.westerndigital.com/dna-data-storage-the-next-chapter/>
- <https://www.scientificamerican.com/article/dna-the-ultimate-data-storage-solution/>
- <https://youtu.be/wxStlzunxCw?si=21i5NPKEdnUJ0q00>
- https://youtu.be/qJMjYLYqbgY?si=fs5YwwGuK8_r_Hi-

Thank you!

What do you think about this presentation ?