BIOMEMORY
Sustainable Storage Solutions

Forum TERATEC 23

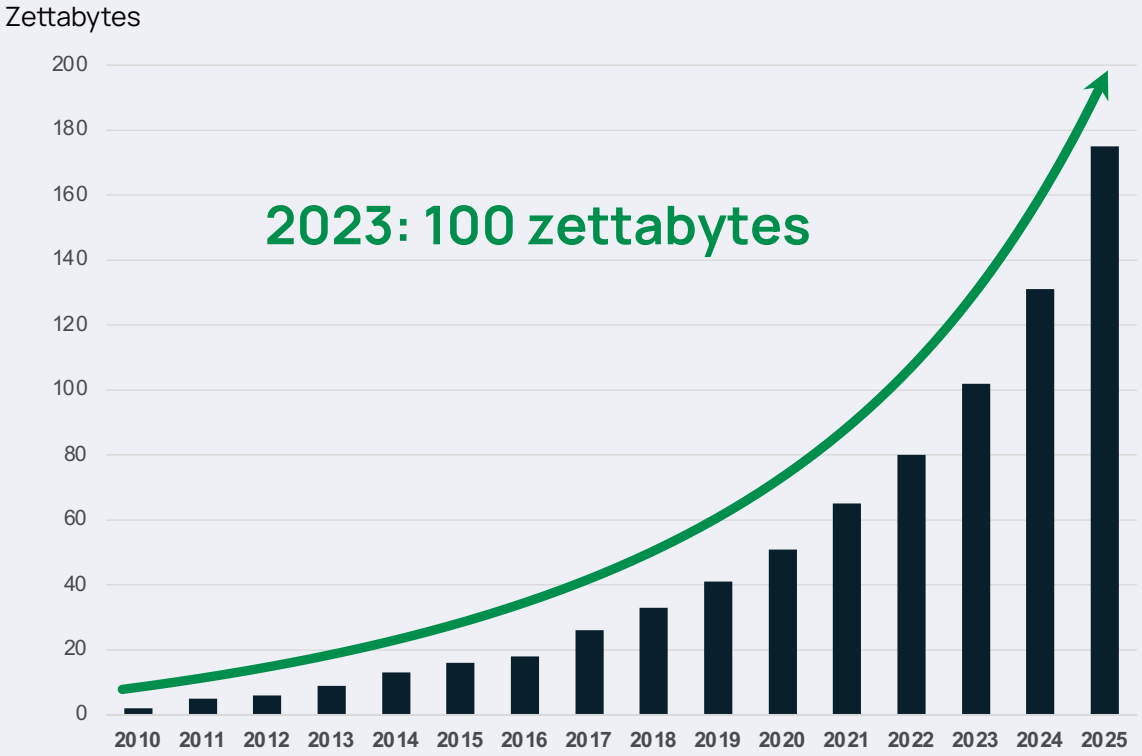
2023 BIOMEMORY

Making DNA **Data Storage** a Reality

Erfane ARWANI, CEO and Co-founder - erfane@biomemory.com

Current digital media are limited and **storage** needs will **explode**

We produce twice as much data every **3 years**



We are less and less able to store the huge amounts of data we produce

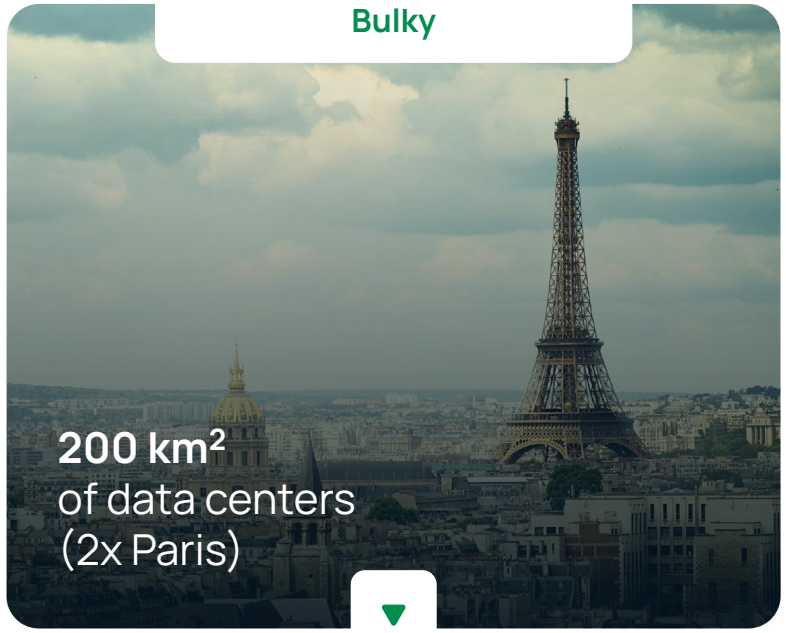
"If today we are capable of storing about 30% of the information we generate, by 2030 we'll be able to store about 3%"

Dr Karin STRAUSS,
Microsoft Research

Microsoft

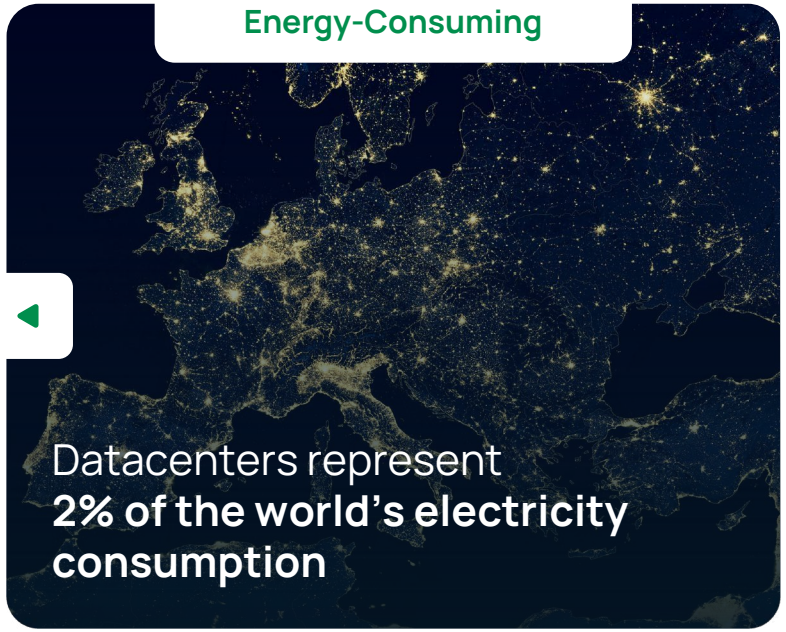
Problems with current digital **media**

Bulky



Fragile

Energy-Consuming

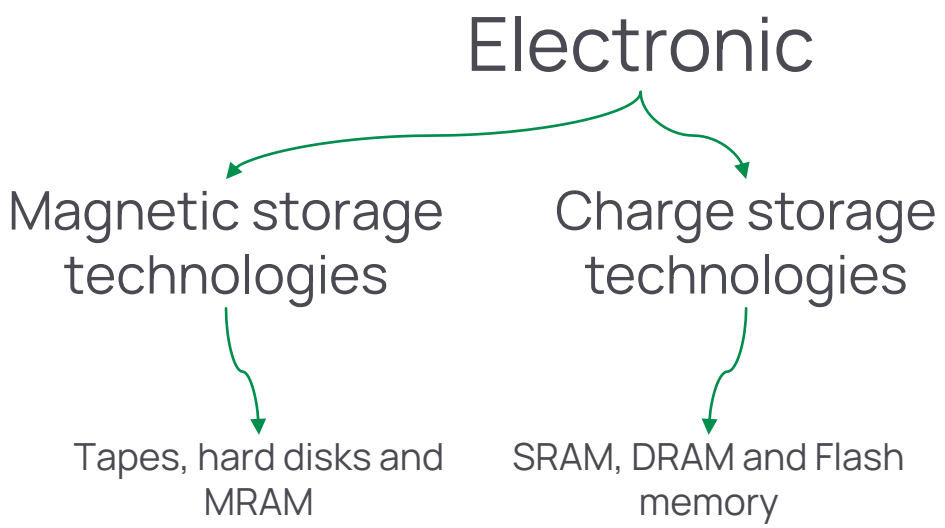


Main storage devices in **datacenters**

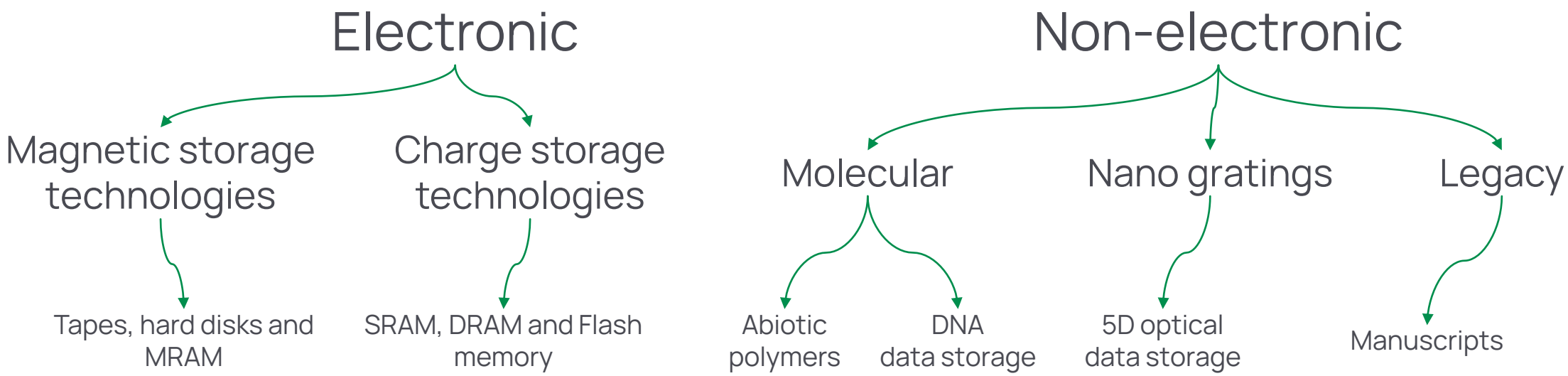
Remastering
the storage devices

- Assess the existing storage infrastructure
- Identify the new storage technology
- Plan the migration
- Data backup and verification
- Data migration
- Testing and validation
- Decommissioning old storage devices
- Monitoring and optimization

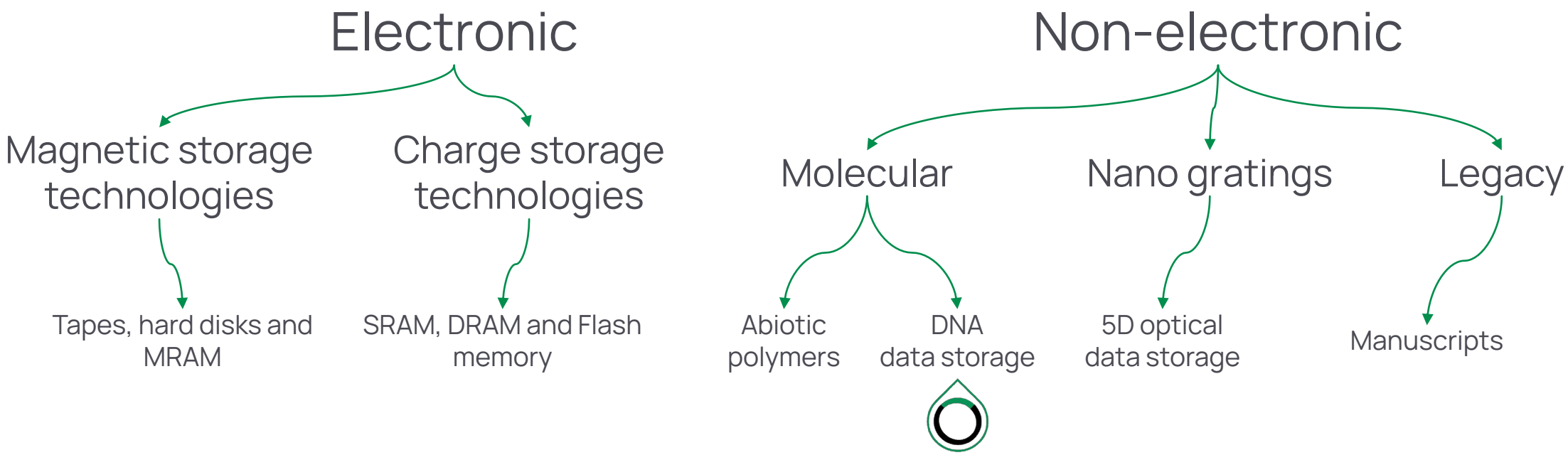
Categories of Data Storage Technologies



Categories of Data Storage Technologies



Categories of Data Storage Technologies

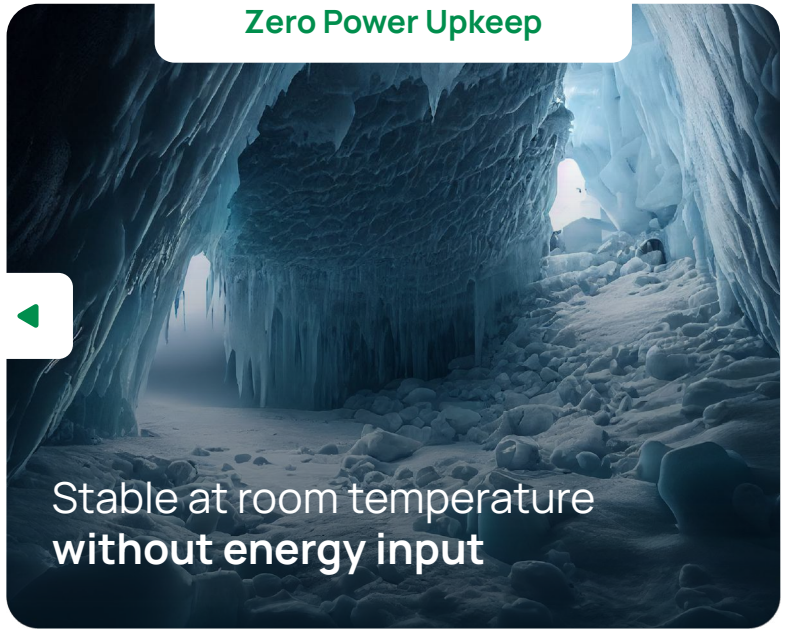


The **DNA** Data Storage

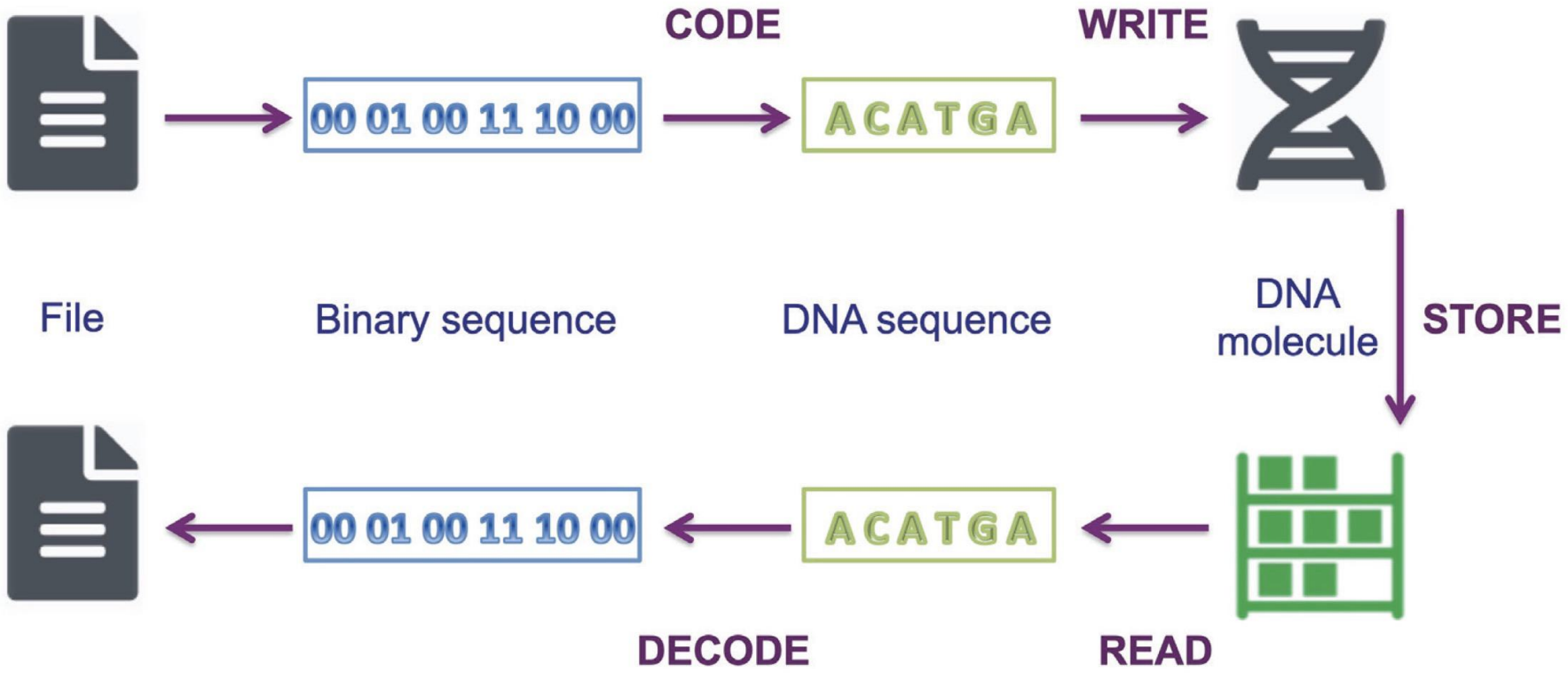
High Storage Density

Longevity & Stability

Zero Power Upkeep



End to end
DNA Data Storage

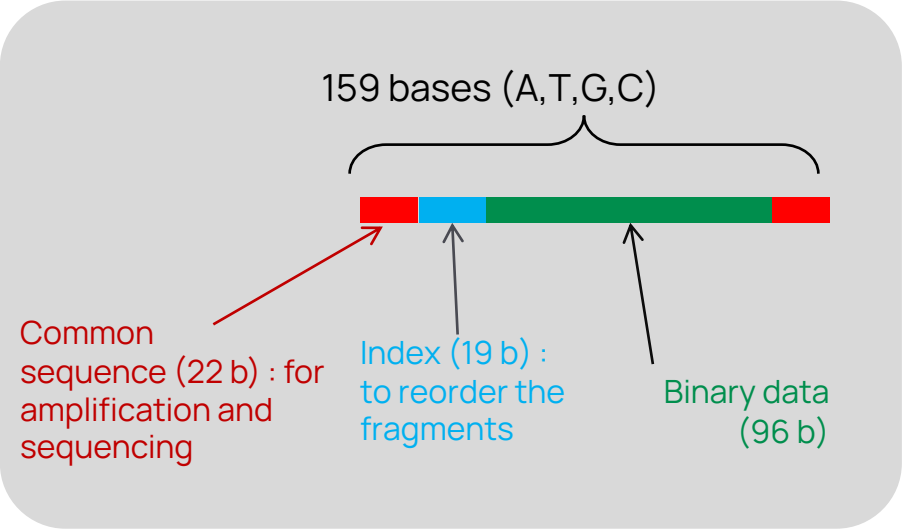


Rapport de l'Académie des technologies. François Képès. Octobre 2020

State of the art
Storage on **oligonucleotides**

Oligonucleotides = short DNA fragments

- Synthesized chemically
- Maximum 200 bases
- Single strand

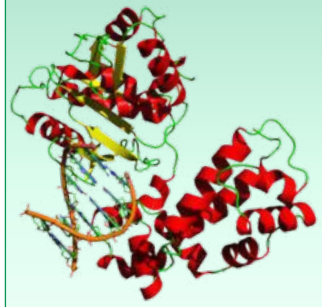


2012, **650 kB** = › **54 898** oligonucleotides — Church *et al.*, 2012 Science

2018, **200 MB** = › **13.4 million** oligonucleotides — Organick *et al.*, 2018 Nature Biotechnol

Microsoft Research · UNIVERSITY of WASHINGTON · TWIST BIOSCIENCE

Need for **more efficient** DNA data storage systems



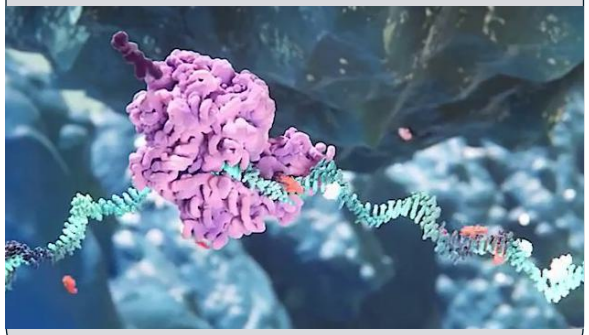
Our vision

- ✓ Exploiting the potential of Nature
- ✓ Employing biological approaches to overcome DNA storage constraints

Images: Science et Avenir, Wikipédia

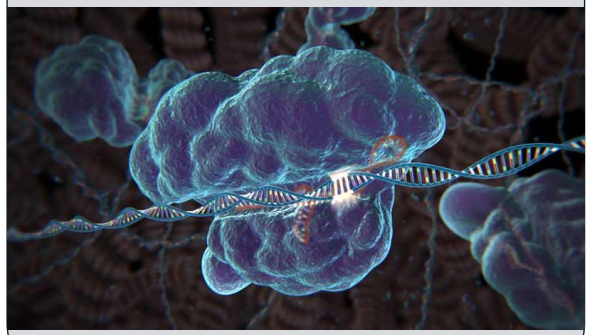
Nature has already figured it out

DNA copy

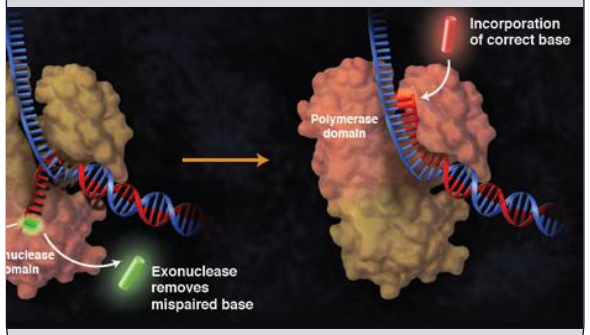


DNA reading

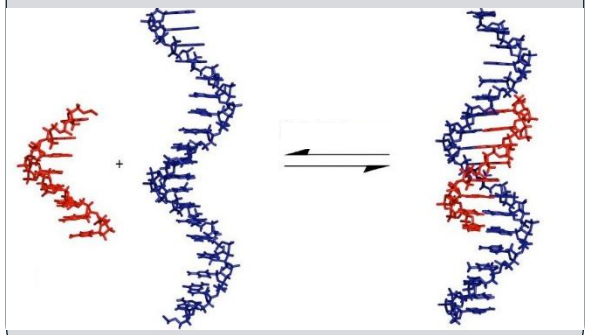
DNA editing



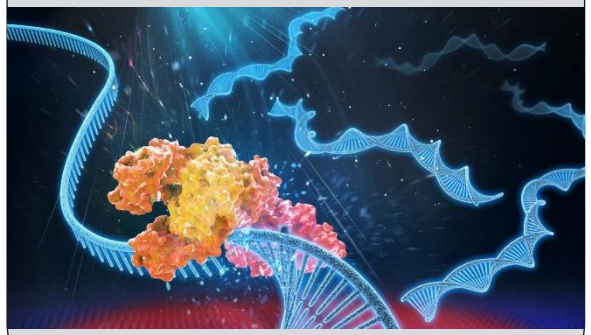
Error correction



Flash random access



Signal amplification



Images : Random42, Neb, Gesundheitsindustrie, Phys

Nature has already figured it out

DNA copy

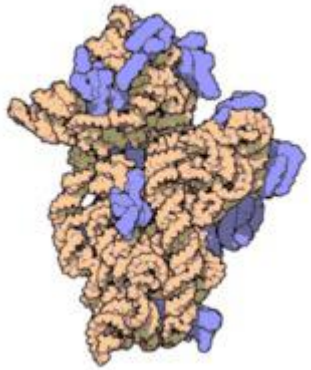
DNA reading

DNA editing

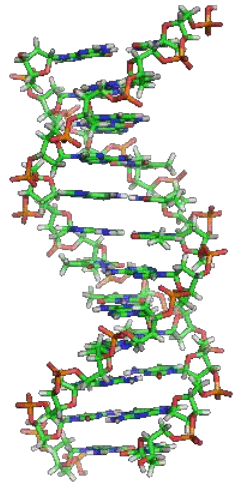
Err

ation

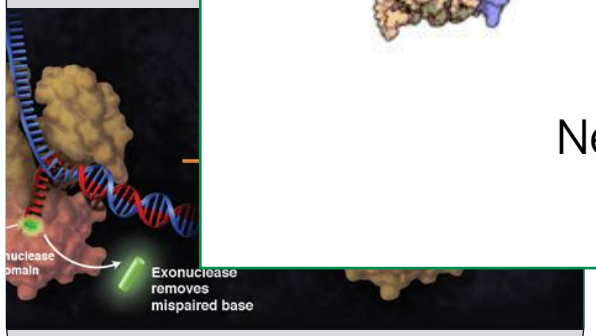
Domestication and **adaptation**



Synthetic biology



New technologies of DNA storage



Images : Random42, Neb, Gesundheitsindustrie, Phys

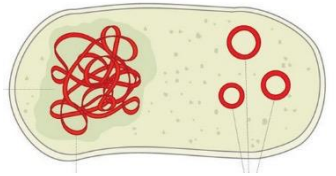
Biomemory strategy
Biocompatible DNA

The mainstream strategy

Oligonucleotides = single strand
- Short length: ≈ 200 bases
- 70 000 fragments /MB

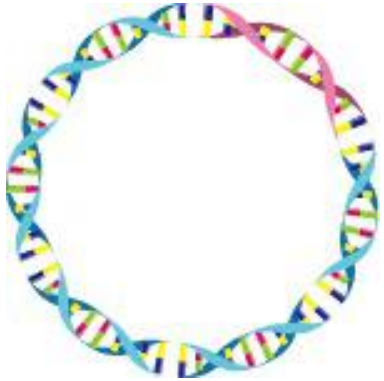
Living organisms

Very Long double-stranded DNA
- Long molecules: kb-MB
- 1 fragment /MB



Chromosome Plasmids

Biomemory



Long circular double-stranded DNA
- Long molecules: 30 kb
- 1000 fragments /MB

→ Biosafe: no biological code, anti-biohacking

→ Copy at low cost and ultra-low error rate

DNA DRIVE
A bio-inspired DNA storage strategy

The mainstream strategy
storage on oligonucleotides

Pool 1

Pool 2

Pool 3

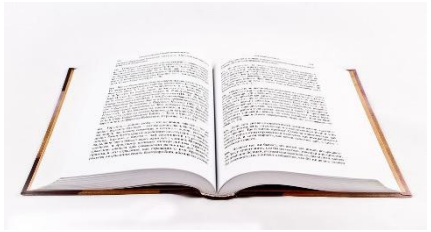
DNA DRIVE
A bio-inspired DNA storage strategy

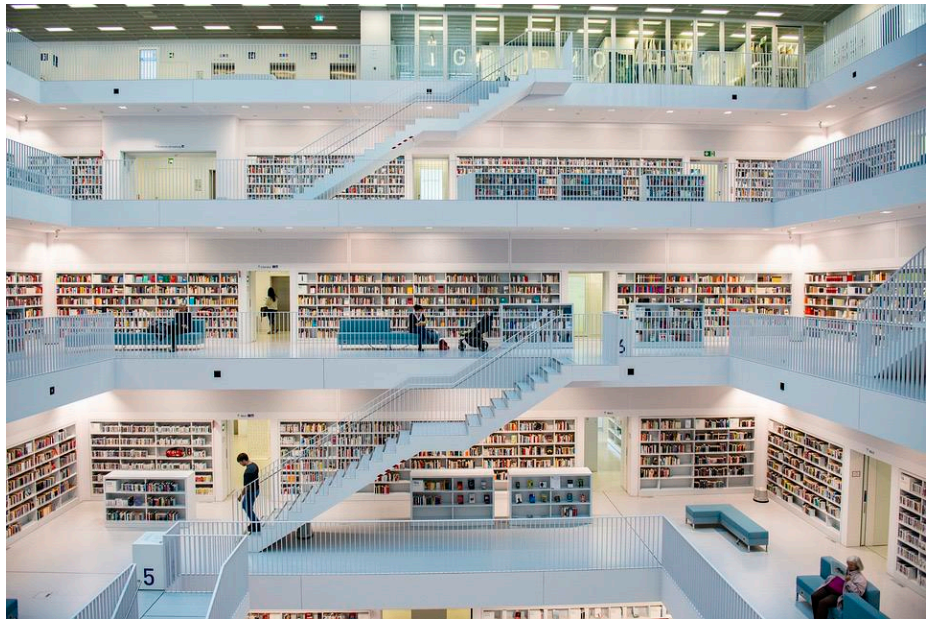
The mainstream strategy
storage on oligonucleotides

Pool 1

Pool 2

Pool 3





DNA DRIVE
A bio-inspired DNA storage strategy

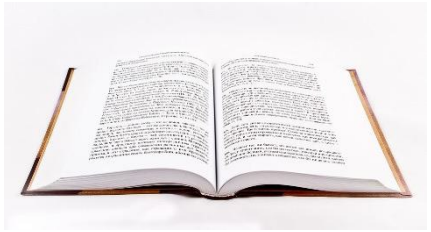
The mainstream strategy
storage on oligonucleotides

Pool 1

Pool 2

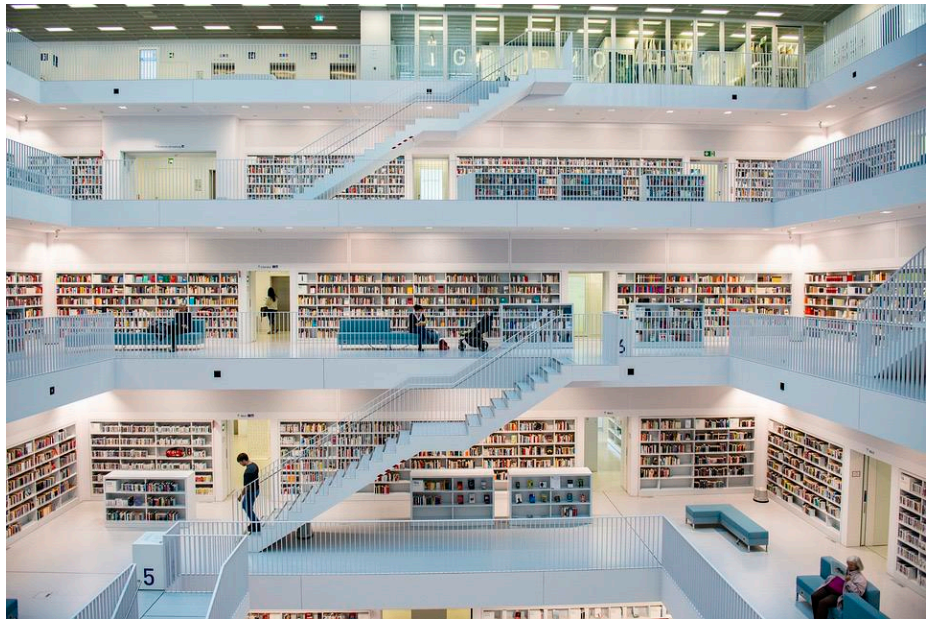
Pool 3

Track



Pool

Array



DNA Drive

DNA DRIVE
A bio-inspired DNA storage strategy

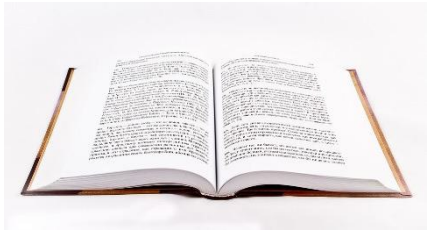
The mainstream strategy
storage on oligonucleotides

Pool 1

Pool 2

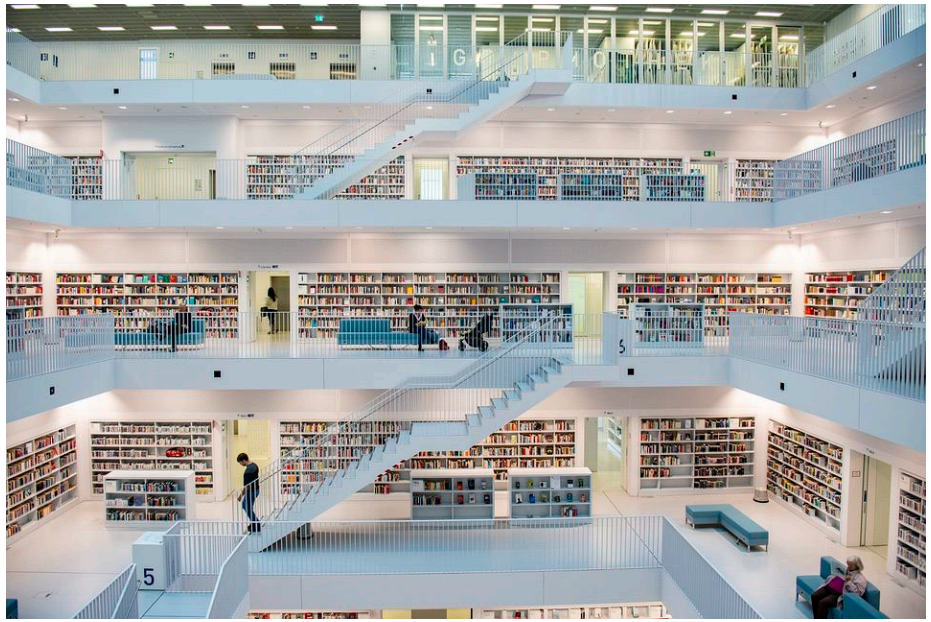
Pool 3

Track



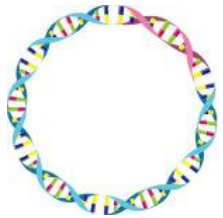
Pool

Array

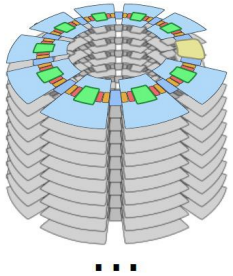


DNA Drive

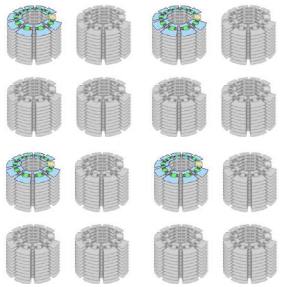
Unlimited number of Arrays



Divided in sectors

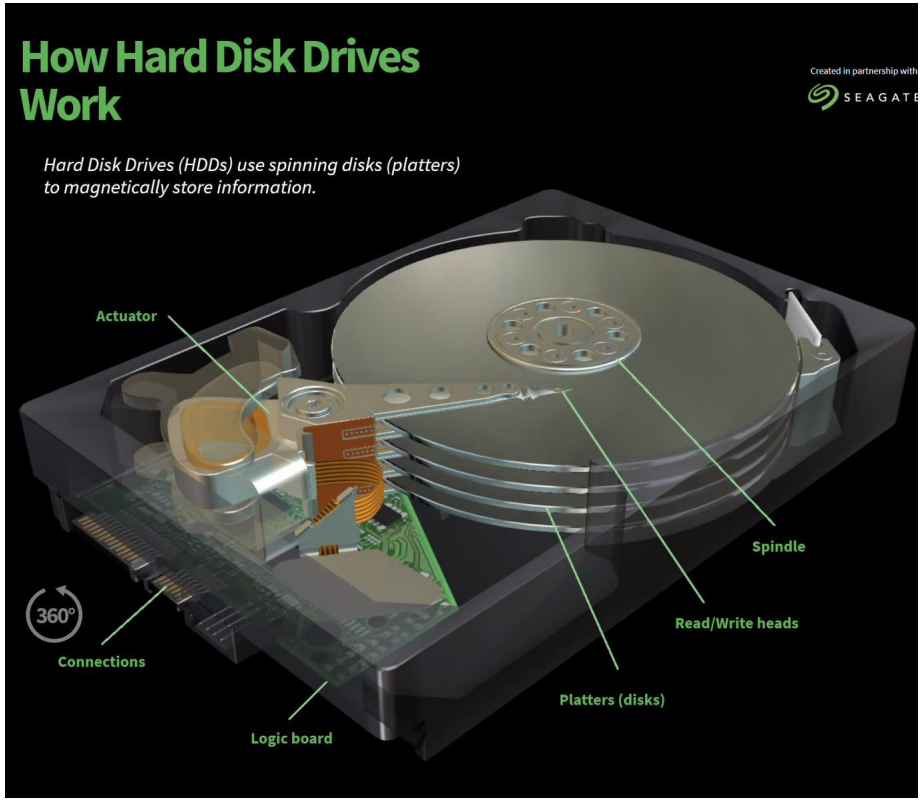


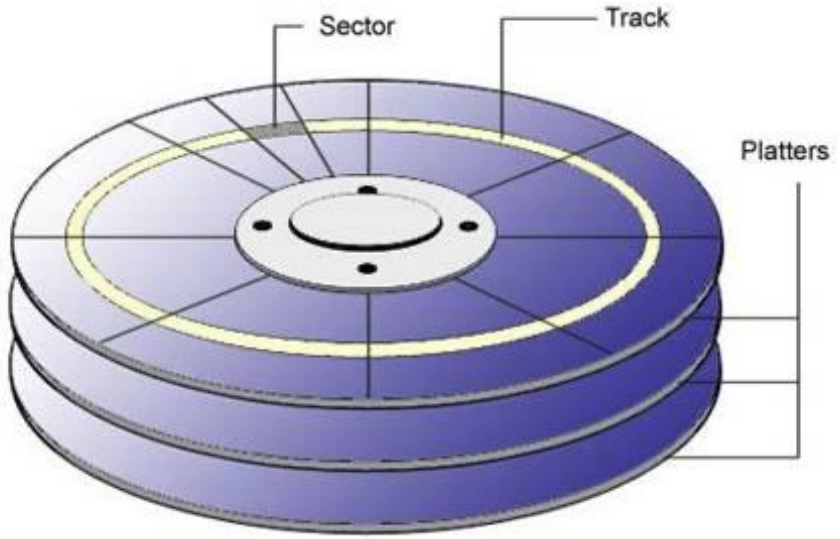
Thousands of Tracks

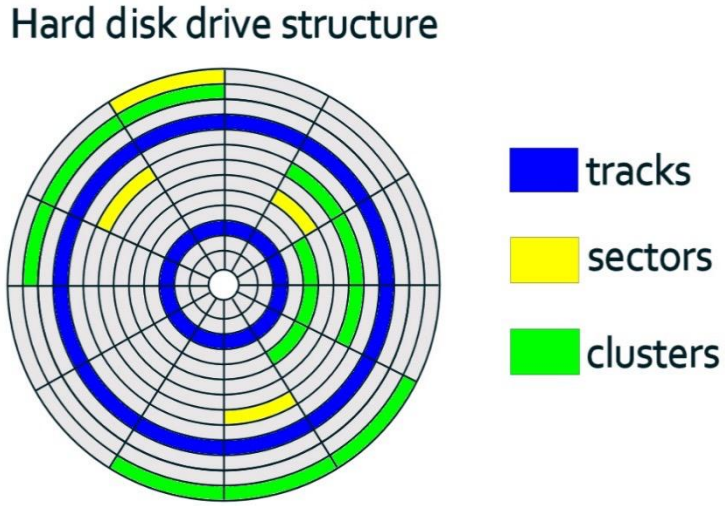


Hundreds of Pools

Hard disk drive
Physical organization

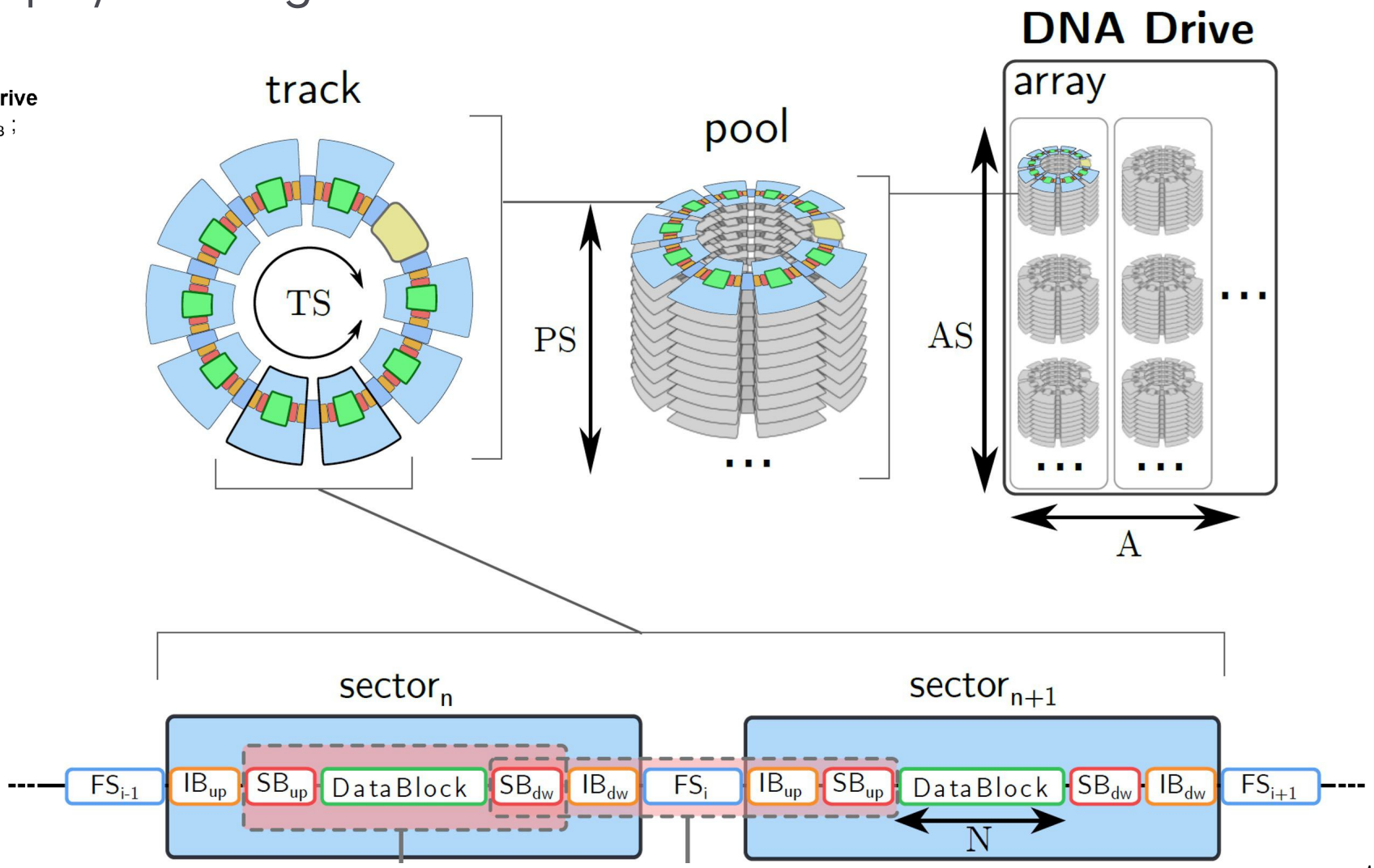






DNA Drive
Multi-scale physical organization

Physical properties of the **DNA Drive**
- **B** Sector size =N_{DB}+N_{SB}+N_{IB} ;
- **TS** Sectors/track ;
- **PS** Track /pool ;
- **AS** Pools/array ;
- **P** total pool number **;**
- **A** total number of arrays



Maes et al. 2022 BioRxiv

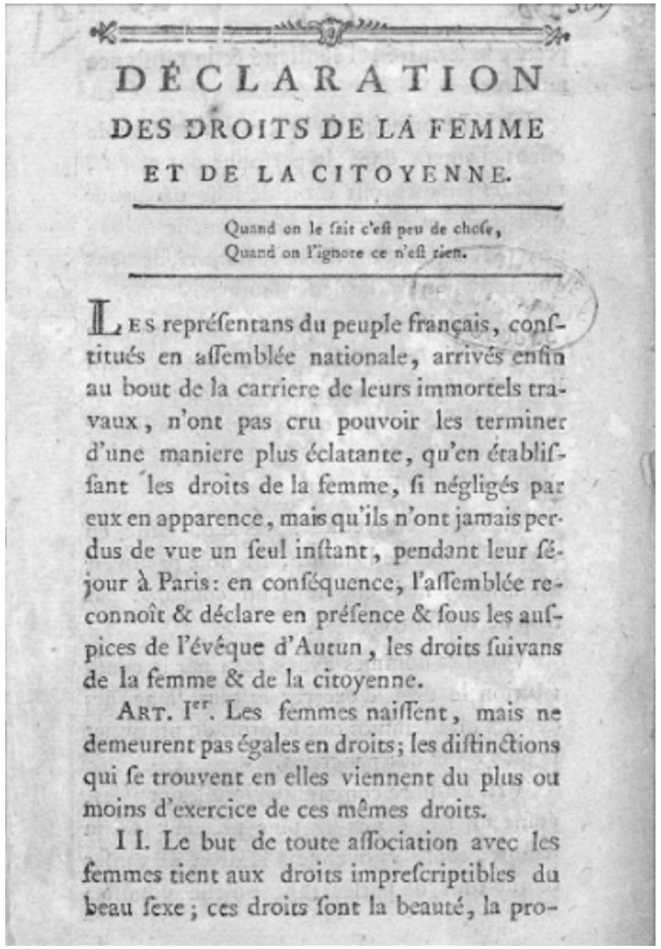
DNA DRIVE
A bio-inspired DNA storage strategy

DNA Drive

The next generation of cold storage media

- Unlimited total capacity
- Low cost copy
  very low copy error rate
- Compatible with all sequencing technologies
- Any binary file system
  organized in physical DNA sectors, tracks, arrays..
- Compression, random access
- Powerful error correction code
- Biosafe
  Non-biohackable by design
- Fully automatable

Proof of concept
historical texts saved for eternity


DÉCLARATION
DES DROITS DE LA FEMME
ET DE LA CITOYENNE.

Quand on le fait c'eſt peu de choſe,
Quand on l'ignore ce n'eſt rien.

LES repréſentans du peuple français, conſtitués en aſſemblée nationale, arrivés enfin au bout de la carriere de leurs immortels travaux, n'ont pas cru pouvoir les terminer d'une maniere plus éclatante, qu'en établiſſant les droits de la femme, ſi négligés par eux en apparence, mais qu'ils n'ont jamais perdus de vue un ſeul inſtant, pendant leur ſéjour à Paris: en conſéquence, l'aſſemblée reconnoît & déclare en préſence & ſous les auſpices de l'évêque d'Autun, les droits ſuivans de la femme & de la citoyenne.

ART. Ier. Les femmes naiſſent, mais ne demeurent pas égales en droits; les diſtinctions qui ſe trouvent en elles viennent du plus ou moins d'exercice de ces mêmes droits.

II. Le but de toute aſſociation avec les femmes tient aux droits impreſcriptibles du beau ſexe; ces droits ſont la beauté, la pro-

Officially stored
at the French National Archives
since November 23rd 2021

ARCHIVES NATIONALES
Fontainebleau - Paris - Pierrefitte-sur-Seine

DNA writing mainstream technique
The chemical synthesis

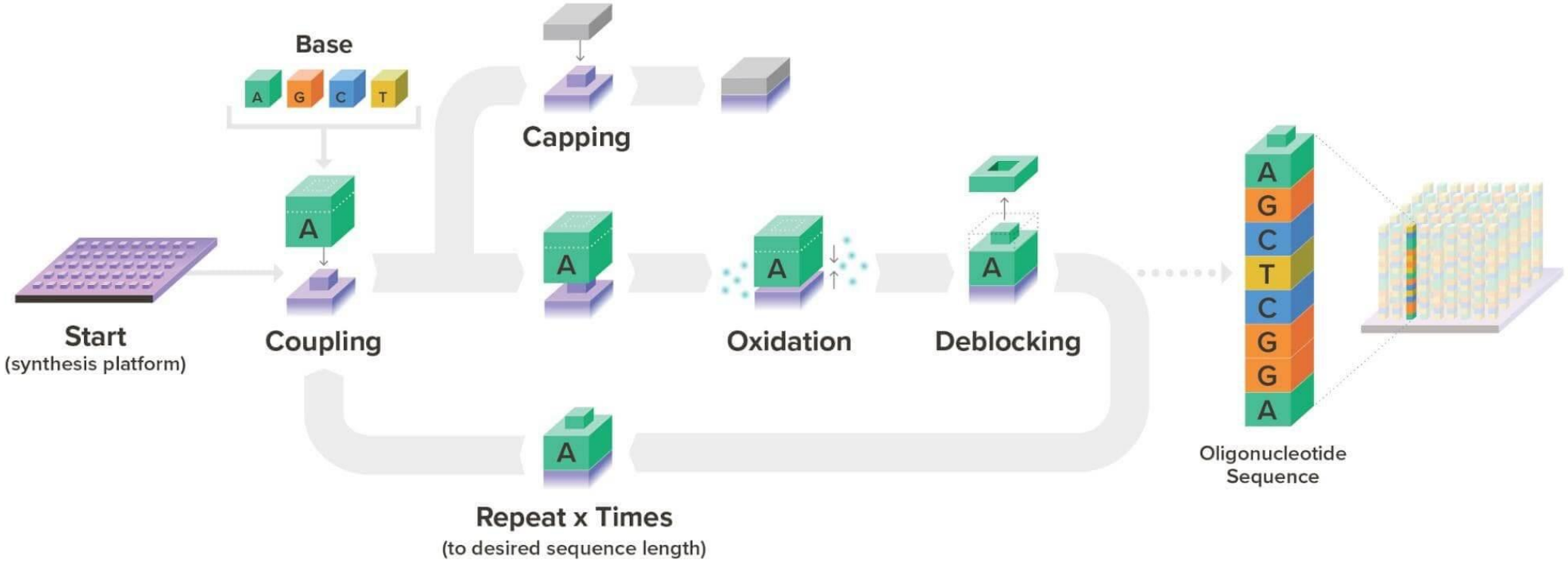
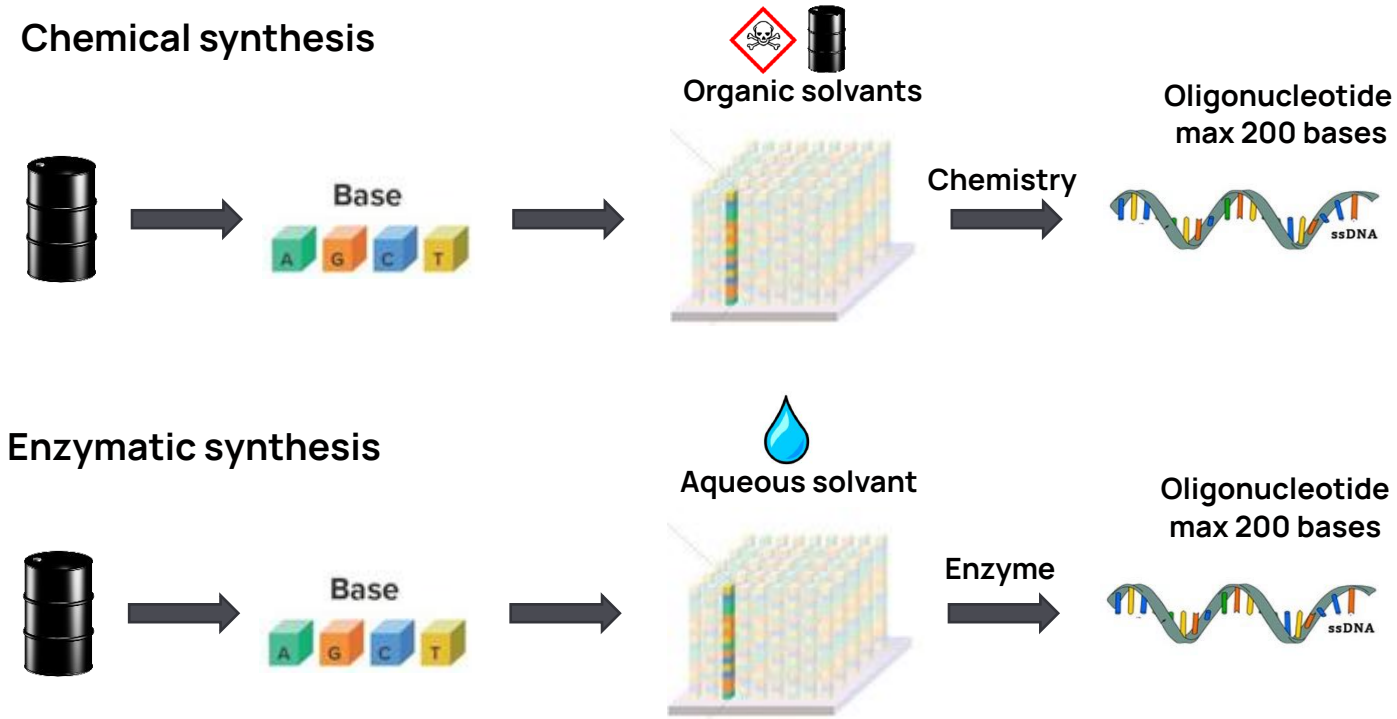


Illustration : Twist Bioscience

Short fragments (<200 nt)
Building blocks from petrochemical phosphoramidites
Utilizes hazardous, flammable organic solvents
Expensive DNA data storage: >$1000/Mo

DNA writing
Synthesis

The mainstream strategy
storage on oligonucleotides



DNA writing
A complexity problem

Biology: you need to be able to synthesize **all** possible DNA sequences

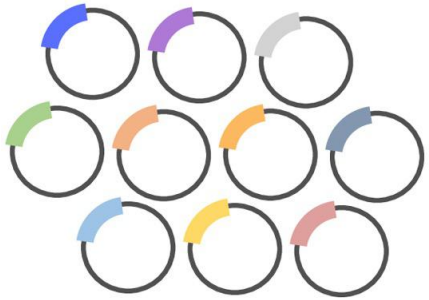
8 nt = › 4^8 = 65,536 sequences

Digital data: *limited* complexity

8 bits = › 2^8 = 256 sequences

Biomemory solution:

A collection of 256 bricks to write any binary file

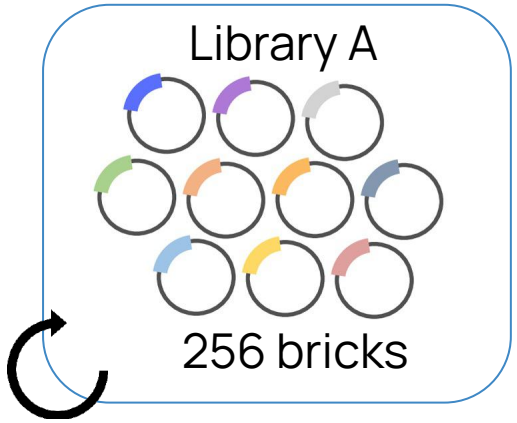


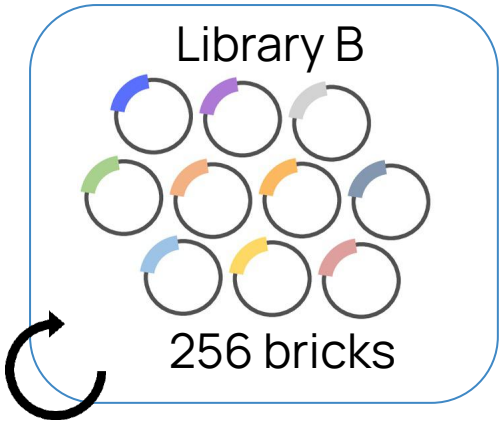
256 Self-replicating DNA bricks

Production at
- High volume
- Low cost
- Low error rate

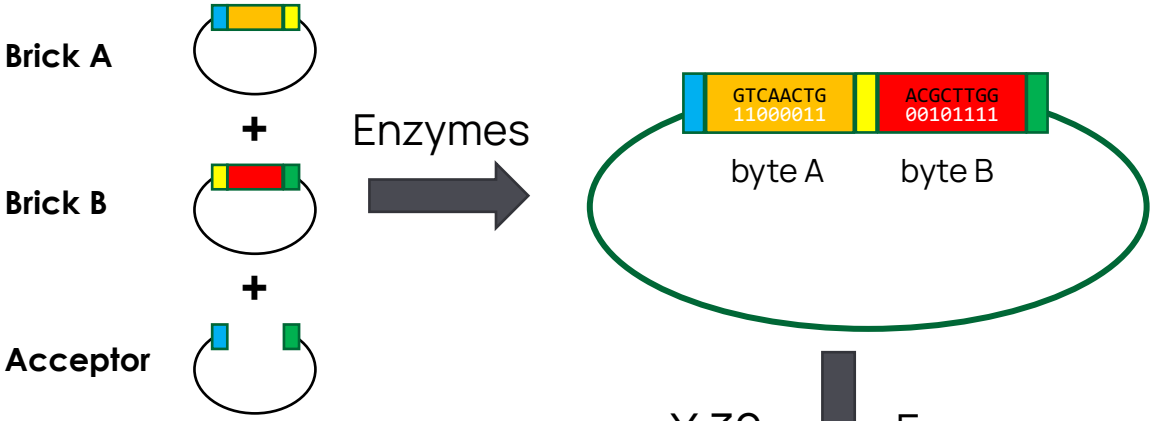
Biodata
Multi-level assembly

Two self-replicating libraries



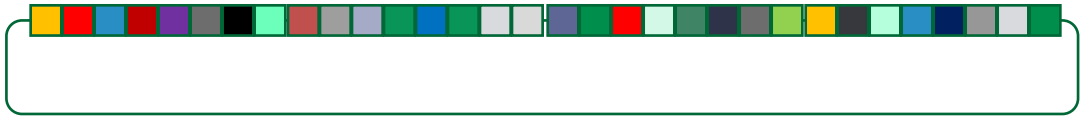


Level 1
2 bytes
Assembly



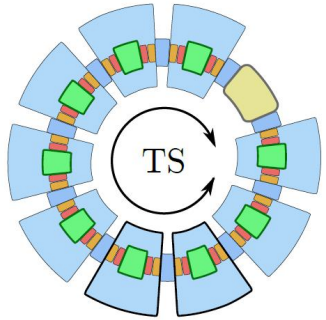
X 32 Enzymes

Level 2
64 bytes
Assembly



X 16 Enzymes

Level 3
1024 bytes
Final assembly

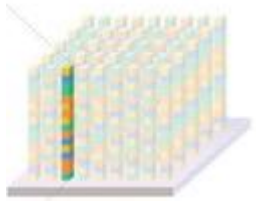


Level 3
Final track

DNA writing
Biomemory's DNA synthesis is radically different

Chemical Synthesis

Organic Solvants

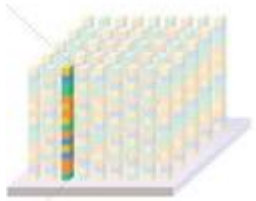


Chemistry

Oligonucleotide
200 bases

Enzymatic Synthesis

Aqueous Solvant



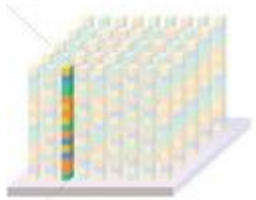
Enzymes

Oligonucleotide
200 bases

Biological Synthesis

Bacteria

Aqueous Solvant



Enzymes

Double stranded DNA
Up to 30k bases

- Potential $1/MB and beyond
- **Renewable** : 100% biosourced
- **Biosafe** : non-biohackable by design

Our **2030** Vision : A rackable DNA Data Storage Server
A business model similar to printers

- ✓Autonomous
- ✓Read/Write
- ✓Exabyte scale
- ✓$1/terabyte
- ✓Removable DNA Drive cartridges
- ✓Removable DNA ink cartridges
- ✓4U rackable server for existing DCs
- ✓No biological expert on site

BIOMEMORY

Sustainable Storage Solutions

Pure Player of
DNA Data Storage

Bio-based
DNA Storage

Let's build the first sustainable data storage solution!

Twitter @BIOMEMORYLABS

Erfane ARWANI, CEO and Co-founder - erfane@biomemory.com