DNA Data Storage Alliance

2024 Review and 2025 Plans

Presented by Dave Landsman (dave.landsman@wdc.com)

DNA Data Storage Alliance

~40 member organizations

Mission

- Create an interoperable storage ecosystem based on DNA as a data storage and compute medium

Scope

- Educate the market to create awareness and adoption of DNA data storage and compute
- Influence and drive R&D and funding
- Develop standards and specifications to encourage ecosystem evolution

SNIA

DNA Data Storage Alliance - Organization

- Board
  - Dave Landsman – Western Digital (co-chair)
  - Esther Singer – Twist Bioscience (co-chair)
  - Stephane Lemaire – Biomemory
  - Marthe Colotte – Imagene
  - Julien Muzard – Entegris
  - David Turek - Catalog
- TWG Chairs
  - Dave Landsman and Esther Singer
- TWG Subgroup, Taskforce, SIG Chairs
  - Data Retention - Dave Landsman
  - Codecs - Manish Gupta
  - Interoperable Interfaces – Shruti Sethi
  - Biosecurity - Esther Singer & David Turek
  - Roadmap - John Hoffman

SNIA

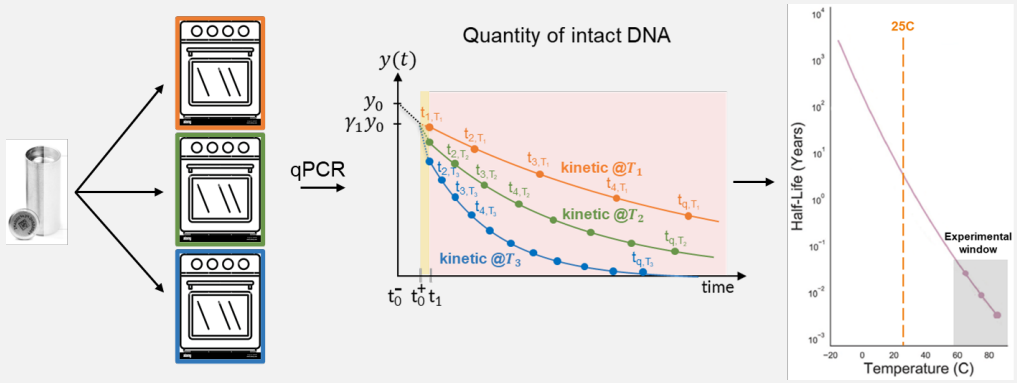
DNA Data Storage Alliance - 2024 Accomplishments

- Specs and Publications
  - Alliance's 2nd specification: DNA Stability Evaluation Method for DNA Data Storage Containment Systems v1.0
  - Chapter on DNA data storage: IEEE Mass Storage Roadmap Update
- Presentations
  - FMS
    - DNA Data Storage – Landsman
    - End-to-End DNA Data Storage System Concept
  - SDC
    - End-to-End DNA Data Storage System Concept (video)
  - Storage Technology Showcase
    - DNA Data Storage Alliance Overview
    - DNA Data Storage Alliance Technical Roadmap
  - Library of Congress Designing Storage Architectures Forum
    - Data Retention Metrics in a DNA Storage System
  - Other
    - Podcast (SNIA Experts on Data): DNA, The Future of Data Storage
    - Monthly all member meetings

SNIA

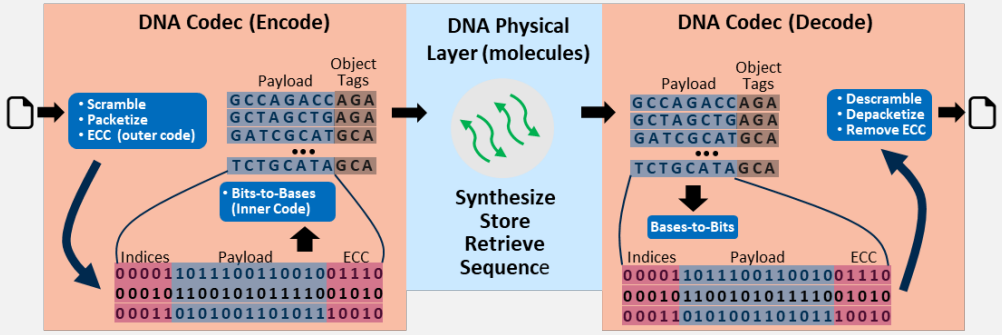
DNA Data Storage Alliance - 2024 Accomplishments (2)

1) Data Retention TWG Subgroup



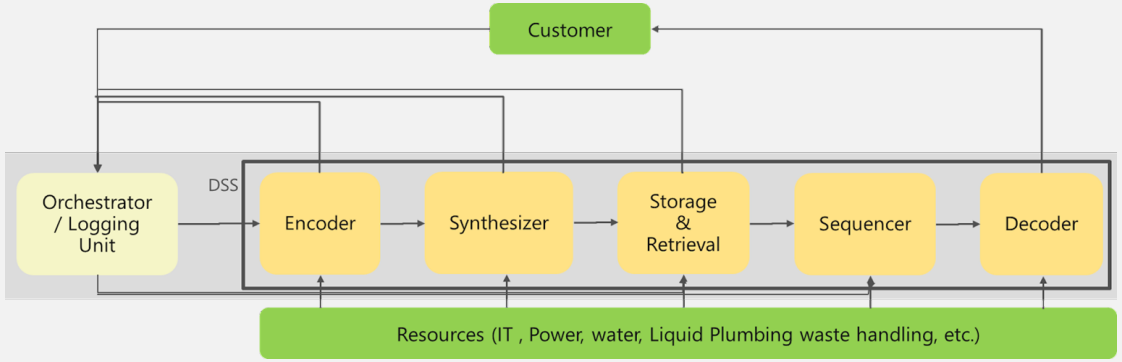
- **V1.0 Published**

2) Codecs TWG Subgroup



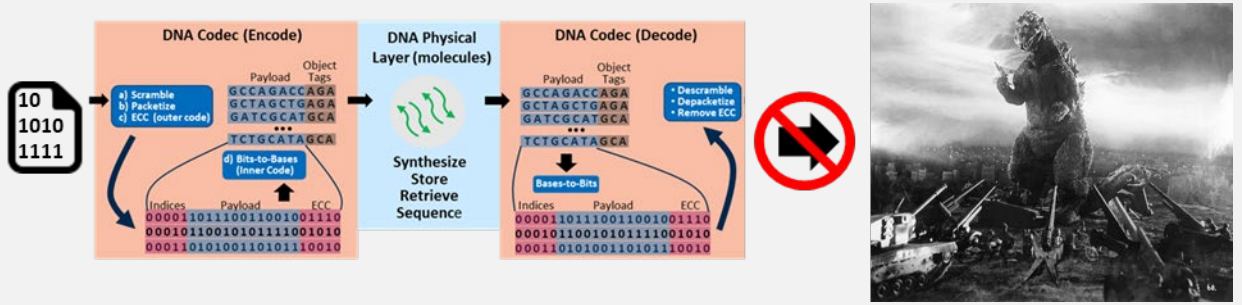
- **Working on "Codec Requirements" spec integration**
- **Open source codec TBD**

3) Interoperable Interfaces TWG Subgroup



- **Working on spec integration**

4) Biosecurity Task Force



- **Developing Alliance position for data storage use case**
- **Evaluating "mathematical proof" vs. "vendor trust" model**

SNIA

DNA Data Storage Alliance - 2024 Accomplishments (3)

- Satellite workshop at ISIT 2024: Coding Theory and Algorithms for DNA-based Data Storage

- Topics of Interest
  - Sequence reconstruction and codes correcting edit errors
  - Emerging sequencing technologies
  - Coding for native DNA-based data storage systems
  - Information aspects of high throughput synthetic biology
  - Machine learning approaches for data reconstruction in DNA-based storage systems

- **60 attendees; 9 plenary talks; 16 posters**
- **IEEE JSAIT special issue to follow in late 2025 (see next slide)**
- **Interesting topics for further review came out of the workshop discussions**

- Speakers
  - Robert Grass (ETH Zurich) - In-depth analysis of the error sources of DNA data storage routines: challenges for error correction codes
  - Olgica Milenkovic (University of Illinois Urbana-Champaign) - DNA Tails for Molecular Flash Memory, Watermarking and Metadata Storage
  - Daniel Bedau (WD Research) - Solid State Nanopores – A Channel with Unique Properties
  - Zohar Yakhini (Technion / Reichman University) - Information and data science challenges in using synthetic nucleic acids
  - James Diggans (Twist Bioscience) - Apportioning error: all current sequencing options are bad
  - Daniella Bar-Lev (Technion) - Universal Framework for Parametric Constrained Coding
  - Natalio Krasnogor (Newcastle University) - Error Correcting Strategies for Storing Short (<1KB) DNA Barcodes In Vivo
  - Adrian Vidal (Monash University) - Learning the Hidden Markov model of the nanopore sequencing channel
  - Hsin-Po Wang (UC Berkeley) - Geno-Weaving: Low-Complexity Capacity-Achieving Data Storage on DNA

SNIA

DNA Data Storage Alliance – 2025 Plans

- Events and publications
  - Conference: Storage and Computing with DNA 2025, Paris, June 19-21
  - Continue other forums we have presented at, as warranted (FMS, SDC, …)
  - IEEE JSAIT special issue: *Information and Coding Theory Aspects of DNA-based Data Storage*
- Finish ongoing work
  - Technology Roadmap/Whitepaper
  - Codec Requirements
  - End-to-End Interoperable Interfaces
- New work
  - Biosecurity: Position on regulatory scope (not overreach) needed for DNA data storage
  - Data retention: Calculator for what is “enough” DNA left to guarantee data recovery?
  - Coding: Characterizing solid state nanopore channel
  - Archive self-discovery: Possible alternative to Rosetta
  - Always looking for other possible topics

SNIA

Why get involved?

- Industry Impact
  - First alliance in this new field; shaping industry as it's being built
- Segment Relevance
  - The storage hierarchy needs a new layer for zettabyte scale storage
- Why join?
  - Fascinating multi-disciplinary field: storage HW/SW, materials, semiconductors, chemistry, biology
  - Opportunity to be part of a birth of a new technology for storage
- Contacts
  - Dave Landsman (dave.landsman@wdc.com)
  - Esther Singer (esinger@twistbioscience.com)

SNIA