EE 491: Senior Design

sddec23-05

DNA Data Storage

Jan. 2023 – Dec. 2023
Client: Professor Meng Lu
Faculty Advisor: Professor Meng Lu

Team Members

Astha Upadhyay, Caden Wortman, Colin Frank, Evan Walters, Anna Hackbarth

Outline:

- Problem Statement (1 Person)
    - Brief intro of your project
        - Problem – Data Storage
        - Why DNA?
            - DNA has a tremendous capacity for data storage
        - How do we take DNA and utilize it for storage?
            - Create DNA from scratch using a 3D printer and the ADS Codex to translate information from binary to biology, put into DNA
- Requirements & Constraints (1 Person)
    - Explain all relevant requirements - functional, environmental, economic, etc.
    - Requirements
        - Low-cost - economic
        - Efficient - economic
        - Sustainable – environmental and functional
    - Constraints
        - DNA tends to mutate
        - There are so many possible combinations
        - Decoding the information is timely
- Engineering Standards (2 people – 1 from hardware, 1 from software)
    - Explain the standards relevant to your project focus - circuit standards, digital design standards, software development standards, etc.
    - Hardware Standards
        - 3D Printer
        - Flow Control
    - Software Standards
        - ADS Codex

        - Fusion 360
- Intended Users and Uses (1 Person)
    - Use-Case diagrams, personas, etc.
    - Who is going to use this technology now? 5 years from now? 10 years from now?
    - What are the long-term goals of this project?
    - Use-case and persona review

Introduction/Problem Statement (1st Person – Astha)

In our modern world we have an ever-growing need for data storage, protection, and privacy. Countless people have their personal information stolen every day and we as engineers are tasked with finding a solution to not only our information security problem, but also the data storage problem. The solution, as cheesy as it sounds, has been inside us all along. DNA is at the forefront of solving our problems.

According to Kilobaser - when converted to digital media, a human diploid cell can store 1.5 gigabytes of data. Now consider that the human body consists of 100 billion cells, that's a lot of information! One gram of dried DNA can store 455 exabytes of data. How much is that? That is 4.55 * 10^20 bytes of data. So how exactly do we take DNA and utilize it for data storage?

Our project aims to create DNA from scratch and where our bodies' DNA contains information that tells us how to be human - this homemade DNA would have no information telling it how to behave, what to do, etc.! The Los Alamos National Laboratory has developed a key technology for DNA storage. The Adaptive DNA Storage Codex (ADS Codex) translates data from the binary language of zeroes and ones that computers understand into the four-letter code biology understands. That information is plugged into the DNA and we're off to the races.

Requirements/Constraints (2nd Person – Caden)

This is where we get into the constraints of our project. Sure, DNA sounds like this amazing tool, but if it were really that simple – why haven't we seen mainstream implementations? There are a few issues that stand in the way. One large requirement that DNA data storage meets relatively easily is sustainability. Is this practice or process sustainable and beneficial for future use? All signs point to yes because DNA has long been known to have incredibly storage capacities and creating DNA from scratch helps avoid ethical dilemmas of standard DNA manipulation.

Liquid Crystal Display (LCD) is an important aspect to our project because that is what displays our image of DNA and what binds the nucleotides to the DNA within the microarray. The current problem with the LCD is that there is light leakage. When it is supposed to be displaying black (turned off) there is light still leaking through the LCD. We need to implement a cooling aspect to the project because it will create a dark spot on the display then any picture taken on the display will be affected by the dark spot.

The previous group for our project had working code which we now have access to, however they did not ever integrate their pieces together to make them easily usable. We are going to need to interconnect the micro controller to the fluid system and create an interface that we can use both systems from, which will be done through the front end that we are going to create. When we finish the front end, we should be able to control the system without having to manually run specific code and increase the overall useability.

Engineering Standards (3rd and 4th Person – Evan and NAME)

Many Engineering Standards can apply to our project, however, there are a few that stand out more than others.

Systems and Software Engineering – Life Cycle Management Part: 1, Guidelines for Lifecycle Managment is one that is quite prevalent within our project. Since we are continuing an existing but incomplete project, Life Cycle Management is very important. We need to review the contents of the project that the previous group left behind and continue down the same path. The 3D models, development stages, and processes of the project are extremely beneficial when documented, so we will continue these practices to upkeep the Life Cycle of the project.

Systems and Software Engineering – Software Testing Part: 2, Test Processes seems to be an obvious choice with the layout of our project. Since we need to integrate the existing micro-fluidic programming with our own, software testing will be very important. This Engineering Standard discusses using multiple generic testing processes to assist in the creation of complicated software. This will help us with troubleshooting software errors and ultimately save us time.

Systems and Software Engineering – Life Cycle Management Part: 6, Systems and Software Integration describes the integration between software and hardware. This applies to our project because we have separate teams working in both sections. This Engineering Standard provides us with guidelines on how to approach the cooperation and integration between the hardware and software sides of the problems.

Users and Intended Use (5th Person – Colin)

Now come some of the most important questions for any design project: who is going to use this and why? Our team believes that there are stages to our project, so our users and clientele base will shift and grow. Starting out now, due to some constraints we listed earlier – our user base is limited to entities with a vested interest in DNA storage. Vested interest is a legal term meaning that you have a personal attachment or interest in a project or property. Entities with a stake in DNA data storage could be research facilities or educational institutions like Iowa State or MIT. They could also be corporations like Microsoft that have a tremendous amount of data that needs to be stored and are concerned about the limited capacity we currently have for data storage.

Down the line, however, our team believes that as DNA data storage becomes a quicker, more efficient and less costly practice – we will see this being implemented by a wider consumer base. We could see within five to ten years that DNA-based storage becomes available for general consumer purchase. Technology evolves so fast and we are hopeful that within our lifetime we can solve the storage problem. We would like to review our persona client who is a researcher interested in furthering DNA based storage. He is an example of a current client or user.

RESOURCES

https://kilobaser.com/dna-data-storage/

https://www.genome.gov/genetics-glossary/Diploid#:~:text=A%20diploid%20cell%20has%20two,so%2046%20chromosomes%20in%20total

https://arstechnica.com/science/2022/09/why-are-hard-drive-companies-investing-in-dna-data-storage/#:~:text=Typically%2C%20each%20base%20of%20the,store%200010%2C%20and%20so%20on.

https://www.scientificamerican.com/article/dna-the-ultimate-data-storage-solution/

https://www.prescouter.com/2017/08/dna-solve-data-storage/#:~:text=The%20main%20challenges%20with%20DNA,about%20400%20bytes%20per%20second.