**Storage of digital data into synthetic DNA**

Rapid technological advances and the increasing use of social media has caused a tremendous increase in the generation of digital data, a fact that imposes nowadays a great challenge for the field of digital data storage due to the short-term reliability of conventional storage devices. Hard disks, flash, tape or even optical storage have a durability of 5 to 20 years while running data centers also require huge amounts of energy. An alternative to hard drives is the use of DNA, which is life's information-storage material, as a means of digital data storage. Recent works have proven that storing digital data into DNA is not only feasible but also very promising as the DNA's biological properties allow the storage of a great amount of information into an extraordinary small volume for centuries or even longer with no loss of information. One of the main drawbacks of DNA data storage is the high cost of DNA synthesis, thus, it is important to control this cost by optimally compressing the input data. We have developed an extended end-to-end storage workflow specifically designed for the efficient storage of images onto synthetic DNA and the feasibility of the process was tested in a wet-lab experiment.

The main goal of this internship is the optimization of our encoding solution using machine learning techniques, increasing the speed of the processes while also allowing to handle large datasets.

**Host Institution:**

The MediaCoding group belongs to SIS Team of the I3S lab, which is the largest information and communication science laboratory in French Riviera. MediaCoding is a group of experts working on image and video coding as well as geometric processing and compression of static surface meshes and animations. MediaCoding researchers are also interested in the analysis of the information contained by the neural code in the visual system, with bio-inspired applications in image and video compression. More recently, they lead activities on the storage of digital images onto synthetic DNA.

**Preferred Qualifications:**

Programming Skills in MATLAB and Python.

**Related bibliography:**


- M. Dimopoulou, M. Antonini, P. Barbry, R. Appuswamy, "A biologically constrained solution for long-term storage of images onto synthetic", *EUSIPCO,* 2019.
- M. Dimopoulou, M. Antonini, "Image storage in DNA using Vector Quantization", *EUSIPCO*, 2020
- M. Dimopoulou, E. Gil San Antonio, M. Antonini, "An efficient sequencing noise resistant mapping for the encoding of images onto synthetic DNA", *MMSP*, 2020


For any further inquiry please contact Dr. Marc Antonini at am@i3s.unice.fr.