

Biologie & informatique : mariage heureux

Sous l'égide d'Université Côte d'Azur et du CNRS, Melpomeni Dimopoulou, 29 ans, Docteur en ingénierie informatique d'origine grecque, travaille au sein du laboratoire I3S sur des sujets passionnants ayant trait à la biologie et à l'informatique, ouvrant un extraordinaire champ de possibilités, comme celle du stockage de données numériques... dans de l'ADN synthétique.

Arrivée en France en 2016, Melpomeni Dimopoulou s'engage dans un Master anglophone de Computational Biology and Biomedicine. « J'ai toujours été passionnée par la biologie et je voulais la combiner avec l'informatique ». Sophia Antipolis est une région attirante pour ces disciplines. « J'ai suivi mon stage au laboratoire I3S (Informatique, Signaux & Systèmes de Sophia Antipolis), en collaboration avec l'équipe de Mediacoding, sous la direction de Marc Antonini, directeur de recherches au laboratoire I3S / Université Côte d'Azur / CNRS ». On lui propose de rester pour une thèse de Doctorat, soutenue en décembre 2020. Elle travaille dans le cadre du projet OligoArchive, financé à hauteur de 3M€ dans le cadre de l'EIC (European Innovation Council). En 2019, Melpomeni remporte la première place régionale de « Ma Thèse en 180 secondes », un concours de vulgarisation scientifique. La jeune surdouée s'était déjà distinguée en proposant le meilleur article lors de la conférence nationale Coresa 2018 puis à la conférence internationale EUSIPCO en 2019. En 2020, elle décroche le financement « Jeune Docteur Innovant Région Sud ».

Son domaine de prédilection ? « Mon rêve était de combiner l'ADN avec le traitement du signal et de l'image ». Des études montrent que l'ADN, qui « stocke » déjà les informations de la vie, serait un excellent support pour accueillir les données numériques. L'enjeu est de taille : les datacenter sont gourmands en espace et en énergie, la data se développe de manière exponentielle et les données numériques stockées auraient une durée de vie n'excédant pas... 20 ans. « En comparaison, on peut toujours décoder, sans erreur, l'ADN d'un mammouth de 40.000 ans ! » sourit Melpomeni.

De l'infiniment petit

Une étude théorique montre que l'on pourrait rassembler l'ensemble des données numériques terrestres sur... 1 g d'ADN ! Et avec les capsules de stockage conçues par la société française Imagene, on pourrait les encapsuler dans l'équivalent d'une boîte à chaussures !

Mais des contraintes existent notamment celles posées par la biologie. Il faut tout d'abord passer du code binaire informatique (01) au code quaternaire A, C, T et G, les sous-unités constitutives de la molécule d'ADN. Mais l'opération de séquençage est sujette aux erreurs...

Avec le soutien de la SATT Sud-Est, une solution, « Paircode » a été brevetée. Les progrès se multiplient et Melpomeni, aidée par l>IDEX UCA^{IEDI}, veut avancer sur ces sujets. Elle a obtenu un financement pour qu'un ingénieur l'accompagne sur une preuve de concept. Dans ce projet, un algorithme va produire un code ADN robuste tout en étant capable de détecter des erreurs dans l'information séquençée, applicable à n'importe quel type de données d'entrée. Avec, à l'horizon 2022, la naissance d'une Startup qui offrira des solutions de stockage longue durée de données comme dans le domaine de la préservation du patrimoine culturel.

Biology and IT: a happy couple

Under the auspices of Université Côte d'Azur and the CNRS, Melpomeni Dimopoulou, a 29-year-old Greek national with a PhD in computer engineering, is working in the I3S laboratory on exciting projects relating to biology and IT that are unfolding a wide range of possibilities, such as storing digital data in synthetic DNA.

When she arrived in France in 2016, Melpomeni Dimopoulou started her master's degree, in English, in computational biology and biomedicine. "I've always had a keen interest for biology, and I wanted to combine it with computer sciences". Sophia Antipolis is very appealing for people working in those fields. "I completed my internship at the I3S laboratory (IT, Signals, and Systems of Sophia Antipolis), working with the Mediacoding team, under the supervision of Marc Antonini, research director for the laboratoire I3S / Université Côte d'Azur / CNRS". She was then offered a PhD, which she defended in December of 2020. She went on to work for the OligoArchive project, funded to the tune of three million euros by the European Innovation Council. In 2019, Melpomeni won a first regional place in the "My thesis in 180 seconds" contest of popular science. This young genius had already gained notice by submitting the best article at the national conference Coresa 2018, and then at the international EUSIPCO conference in 2019. In 2020, she was granted funding reserved for "Young Innovating Doctors of the Southern Region".

What is her favourite field of research? "My dream was to combine DNA with signal and image processing". Studies show that DNA, which already stores life's information, can be an excellent support to save digital data. Data storage is a massive challenge, with datacentres using space and power to store exponentially-growing amounts of information, and with a life expectancy for stored digital data that does not exceed 20 years... "By comparison, the DNA of a 40'000-year-old mammoth can be accurately decoded today!" Melpomeni says with a smile.

Entre nous

Elle aime de son travail...

« ...qu'il trouve une utilité, une concrétisation : la plus belle récompense ! »

Et de la Côte d'Azur...

« ...Antibes, où j'habite et plus généralement l'ensemble de la région : elle rassemble tous les bons côtés de la Grèce, dont, évidemment, la douceur du climat... »

Une sortie...

« ...courir en bord de mer, c'est incomparable ! »

What she likes about her work...

« ...is that her efforts are useful, and tangible: it is the best reward! »

What she likes about the French Riviera...

« ...Antibes, where I live, and more broadly the rest of the region: it has some of the nicer sides of Greece, like its lovely climate... »

An outing...

« ...a seaside run, there is nothing quite like it! »

From the infinitely small...

A theoretical study shows that we could store all the world's digital data on... 1 gram of DNA! With the storage capsules designed by the French company Imagene, we could encapsulate it in the equivalent of a shoebox!

However, there are some limitations imposed by biology. Firstly, the binary code used by computers (0 and 1) must be translated into a quaternary code featuring A, C, T, and G, the subunits that constitute the DNA molecule. But the sequencing operation is subject to errors... With the support of SATT Sud-Est, a solution has been patented under the name "Paircode". Significant advances are being made and Melpomeni, with the support of IDEX UCA^{IEDI}, is making great strides on these issues. She has secured funding to receive the support of an engineer who helps her with proof of concept. In this project, an algorithm will produce a robust DNA code, while also detecting errors in the sequencing of the information, applicable to any type of input data. This should all lead, by 2022, to the creation of a start-up company that will offer long-term data storage solutions and will work to protect our cultural heritage.

MELPOMENI DIMOPOULOU

LABORATOIRE I35 – UNIVERSITE COTE D'AZUR / CNRS

