“AI” for health, misinformation and cultural heritage! Our researchers explain the vast applications of Artificial Intelligence tools! P12-13

EURECOM
Sophia Antipolis

Research MAG

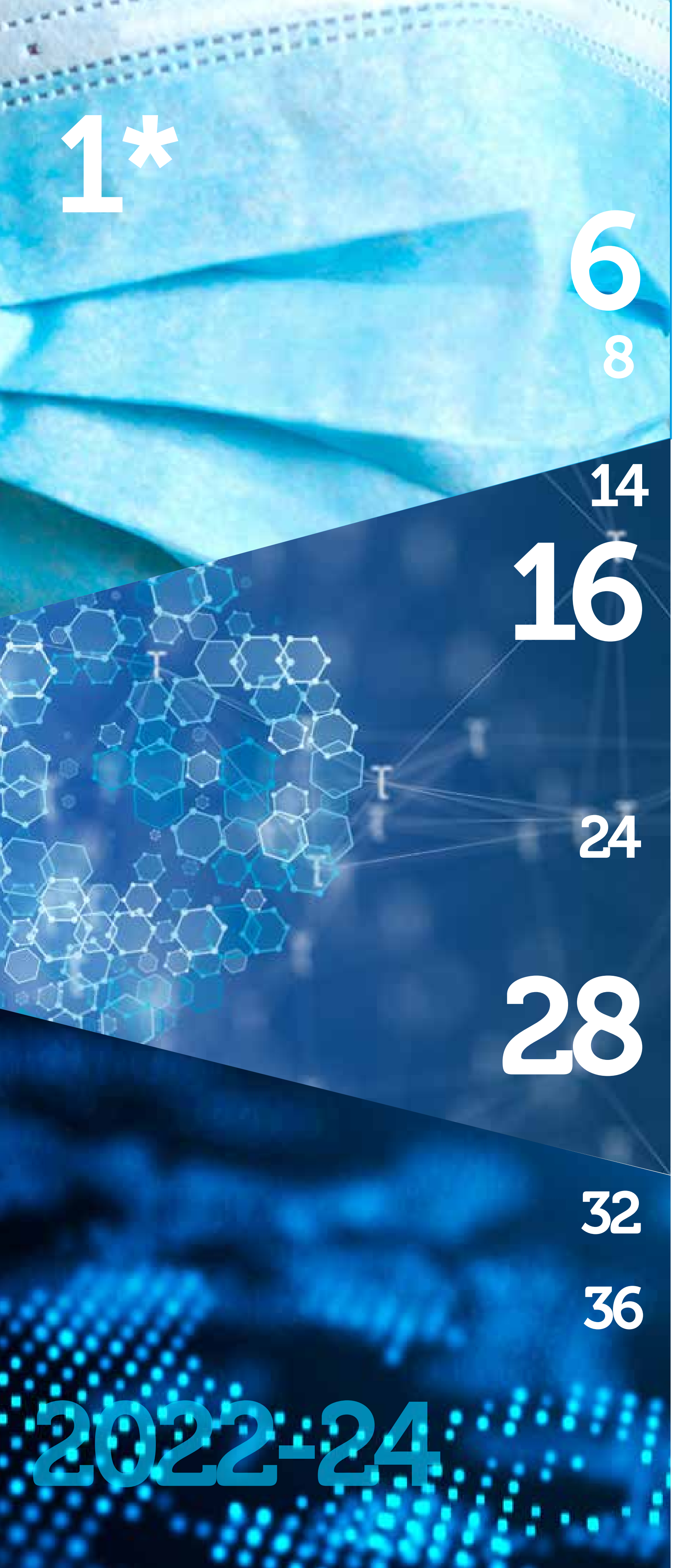
DISCOVER 15 RESEARCH TOPICS THAT WILL TRANSFORM THE FUTURE OF OUR SOCIETY

Automated Cars and Smart Territories: an essential cooperation P34

Drones can ensure our safety & enhance our connectivity. Are we ready to see them fly above our cities? P32

Entering the world of IoT, our devices are vulnerable to cyber attacks! How can we avoid them? P18

Our data can live forever! Learn how DNA can be used to store them! P10



CONTENT

“I would like to congratulate and thank all our teams who, through their commitment, help EURECOM achieve these excellent results!”

David Gesbert
Director of EURECOM

Welcome to EURECOM, one of Europe’s ICT-domain most vibrant research and teaching institutions!

Located in the glamorous French Riviera, at the heart of Sophia Antipolis technology parks, our research mission could ideally be summed up as “excellence with relevance!”. Our labs strive for academic excellence in all the key areas of digital sciences, covering the diversity of topics in cybersecurity, data science for artificial intelligence, networking for the future internet, and next generation mobile systems.

In all these domains, EURECOM is recruiting the brightest minds to push the frontiers of knowledge, competing at the highest international level for grants, citation index, and awards. As an example, no less than 15% of our Faculty have been awarded the highly competitive ERC (European Research Council) grants in the last years, one of Europe’s highest ratio!

Yet, excellence would be lost without much needed relevance. In that respect, tight research collaboration with industry has always lied in the DNA of EURECOM. In fact, several of our industrial partners share with us the governance of our school, steering our research strategy together with 11 of Europe’s best technical universities.

Another advantage of our strong research program lies in the way it positively impacts on the quality of our curriculum, enriching our courses with latest concepts and giving further motivation to our student do well in class and sometimes to engage in research projects with the Faculty.

With our partners, we design tomorrow’s communication and intelligent transport networks, machine learning and cybersecurity solutions and more.

Then, there is societal relevance too. Some of our recent research focuses on use cases ranging from AI for health (automatic cancer detection etc.), energy-aware machine learning, protection of data privacy, green and low-emission wireless networks, and smart autonomous transports to fight congestion in our future smart cities.

All giving great examples for young researchers to realize that computer and mathematical sciences are not just intellectually fun but can also serve the greater good in society.

To all these young (and less young) research enthusiasts, I would like to say happy reading. Take time in discovering our exciting new projects and reach out to us if you care about joining an all-international, English-speaking community of bright minds driven by their passion to invent tomorrow’s digital society!

BIO

Prof. David Gesbert (Fellow, IEEE) is currently serving as Director of EURECOM, Sophia Antipolis, France.

From 1997 to 1999, he was with the Information Systems Laboratory, Stanford University. He was the Founding Engineer of Iospan Wireless Inc., a Stanford spin off pioneering MIMO-OFDM (currently Intel).

Before joining EURECOM in 2004, he was with the Department of Informatics, University of Oslo, as an Adjunct Professor. He has published about 350 articles and 25 patents, 7 of them winning IEEE Best paper awards.

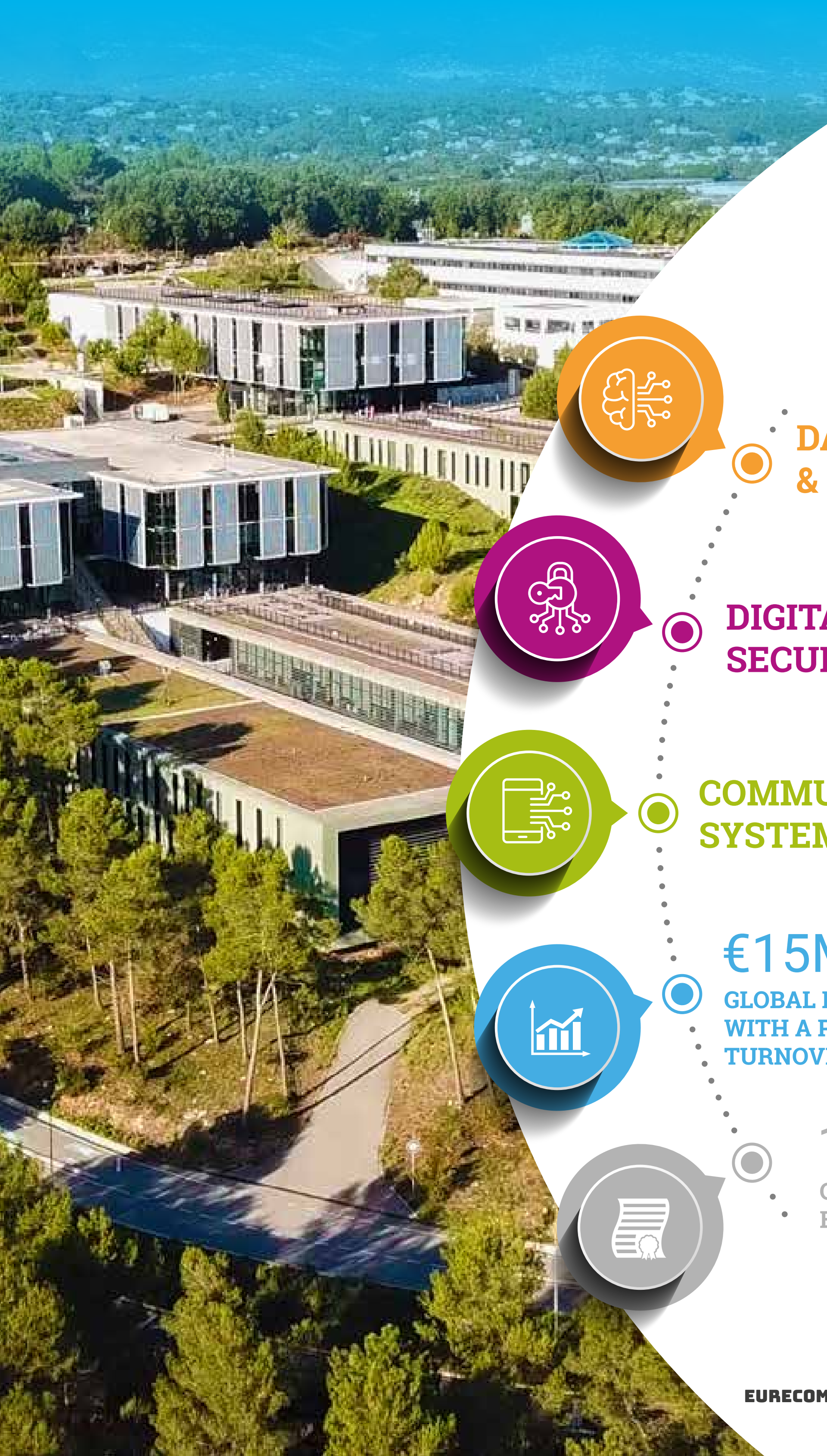
He has been the Technical Program Co-Chair for ICC2017 and has been named a Thomson-Reuters Highly Cited Researchers in computer science.

He is a Board Member for the OpenAirInterface (OAI) Software Alliance. In 2015, he has been awarded an ERC Advanced Grant. In 2020, he was awarded funding by the French Interdisciplinary Institute on Artificial Intelligence for a Chair in the area of AI for the future IoT.

In 2021, he received the Grand Prix in Research from IMT-French Academy of Sciences.

For over 30 years, EURECOM has been a key player in the European research landscape.

Promoting the synergy of industrial and academic cutting-edge research is our approach, in order to address the new scientific, technological and societal challenges of the 21st century.



Our mission is to better understand our world, putting our research expertise at the service of humanity !

DATA SCIENCE & ENGINEERING

DIGITAL SECURITY

COMMUNICATION SYSTEMS

€15M

GLOBAL BUDGET WITH A PROJECT TURNOVER OF €8M

115

CONTRACTS MANAGED EACH YEAR

INSTITUT CARNOT
Télécom & Société numérique

Innovation & transfer towards industry

Opening new paths towards future and emerging technologies, and establishing partnerships with companies to enhance knowledge transfer to tomorrow's industries are two key strategies of EURECOM's policy.

EURECOM supports its students' and researchers' startup projects with the through its incubator program.

26
PROFESSORS

320
MASTER STUDENTS

102
DOCTORAL STUDENTS

51
RESEARCH ENGINEERS + POST-DOCS

12
VISITING SCIENTISTS

Data Science Department

AI at the service of society

The Data Science Department is deeply interested in important questions for humanity such as climate change, sustainability, health, digital preservation. EURECOM is well positioned with respect to these problems, by having meticulously selected faculty members, in order to acquire excellence and expertise in key areas.

Pietro Michiardi
Professor, Head of Data Science Department

Q. Could you briefly present the Data Science Department at EURECOM?

PM. The Data Science department at EURECOM, established in 2016, has, a relatively short history of existence! Over these years, we have meticulously selected our faculty, in order to acquire excellence and expertise in key areas. *I would say that, in our department, the common underlying force is AI and knowledge and information management.*

Under this general umbrella, our research groups are developing strong theory and application domains such as AI for Health, Novel Data Storage, Sustainability, Finance, Cultural Heritage, Information Quality and of course AI for large ICT infrastructures, in collaboration with the other two historical departments of EURECOM.

As a department, we are deeply concerned by the big questions for humanity such as climate change, sustainability, health, digital preservation and we think that it is important for EURECOM to be well positioned with respect to these problems.

Q. Working on diverse application domains in one single department is very interesting and quite rare. Could you describe them in more detail?

PM. A very important domain is AI for health, especially medical imaging, collaborating with hospitals and medical centres, having the opportunity to work with real medical data and practitioners. We are using Machine Learning (ML) models and are able to apply those to poorly annotated medical data and make them robust to noise. We see AI as Augmented Intelligence, involving humans in the loop, providing useful ML tools for medical doctors, in order to help them rather than replace them. There is always a human being involved in the decision process, creating a feedback loop that can help the algorithmic side of ML.

Another important application domain in

our department is Data Storage, where we study new "exotic hardware", for example synthetic DNA, in order to store information, especially related to life sciences, e.g. molecular, genetic data etc. There is an explosive rate at which such type of data is becoming available. At EURECOM, we are working both on the algorithmic and physical aspect of information storage, in order to create the equivalent of a classical database on synthetic DNA. Also, we are working in close relation with bioinformatics teams, to improve the algorithms that are used for the sequencing of genetic data. Sequencing is important for many critical aspects apart from data storage, such as drug discovery, diagnostic tools, biomarkers for rare diseases etc.

Nowadays, quality of information, misinformation and fake news are a real problem with a huge impact in society. In our department, we use techniques from ML and knowledge and information management to fight misinformation in several topics, related to the COVID-19 pandemic, climate change and energy. We are working in collaboration with large companies like Google, but also organisms and authorities fighting misinformation like International-Fact Checking Network (IFCN). In our department, we have a very wide expertise spanning from good theoretical understanding up to implementing online demos that can be used against misinformation. In this domain, we have a capability to cover both theory and practice, which is quite rare.

Finally, we also focus on society-related topics such as AI for cultural heritage preservation. Concretely, we are main partners in two large European consortiums related to a) the olfactory preservation of the past and b) the European Silk-Road preservation. We have strong expertise in Knowledge Graphs and Natural Language Processing and understanding, using it in order to provide insight to our history and culture by using AI technology.

Q. At the same time, there is a very strong theoretical component in the department. What is the main expertise you bring?

PM. We have a group working on the Theoretical aspects of AI, mostly working on the development of mathematical tools for AI and use them for applications in various fields like financial crime investigation, the automotive industry but also sustainability and catastrophic events prevention. For example, we are working with the government of Monaco on an application to study performing simulations about "What if?" scenarios for possible disasters and evacuation plans.

One of our theoretical work, that has attracted quite a lot of attention is the Bayesian modelling of neural networks; we managed to change the perspective of the community on the way Deep Learning is typically done. We have worked to transform historical Deep Learning models, to have the capability to quantify uncertainty, and be able to predict when models are wrong.

Another important topic is representation learning, which is how to learn latent representation of data, that can be used either to encode information in a compact way or to be used as methods to extract important features for any downstream application. As an example, we learn general representations by merging information produced by multiple sensors with applications to self-driving cars that scan the environment using various modalities (cameras, LIDARS, radars, etc...). Concretely, we are providing theoretical tools for AI and ML, that are used for industrial applications with our collaborators like SAP, Huawei, Renault, Oracle.

Q. What is the future perspective of the Data Science Department?

PM. From a strategic point of view, we aim for academic excellence, which we achieve by publishing at top conferences, and by giving talks and seminars to be widely known and recognized in the outside world. From the applications point of view, we need to gain a critical mass of achieved and finished projects, in order to build up our portfolio and increase visibility, especially locally in Sophia Antipolis. We will focus on few flagship applications and position ourselves as experts, maintaining a coherency and establishing our identity for external parties' recognition.

Due to EURECOM's consortium structure (GIE), we are privileged to have top companies and institutions as our members and therefore, facilitate strong collaborations. For example, we would like to extend our partnership with Monaco hospitals, which are top level: here we have the competitive advantage since the government of Monaco is one of EURECOM members.

Also, we have a new member, EDHEC business school and we will be keen to collaborate on a wide range of common research topics. Finally, I am proud to share that we have a recent accepted project, "4DOmics", in collaboration with CNRS, for creating at the premises of EURECOM one of the largest European biobanks with capacity of tens of petabytes. The project's goal is not only to be the archival place for biological data storage, but also eventually, to become the place where algorithms and computations are run.

Q. What are your long-term goals as Head of the department?

PM. I think we have to continue focusing on AI models that are interpretable, computationally efficient and amenable to a precise quantification of uncertainty; these are the three main directions that we are already covering and that we need to continue for sure. Another point is that there is a field attracting a lot of attention recently; the mathematical aspect of statistical physics and ML. Mathematical models used molecular mechanics, for example, are wonderful tools to help us deepen our understanding of how ML actually works and designing new models. In fact, we translate what we know from physics and apply it to ML. We are only at the beginning in the community and I would like to invest towards this direction in the future, at least for the next five years.

Towards AI safely used **for our health**

Maria A. Zuluaga
Professor, Data Science Department

Artificial intelligence (AI), the broad discipline of creating intelligent machines, has been very well established in the everyday technology. But how much can we trust AI for applications that are critical like healthcare, aviation or security? EURECOM's professor Dr. Maria A. Zuluaga, expert in AI and Data Science will guide us through this crucial question.

Daily, we interact with intelligent systems while sending an email, unlocking our phones, using social media, receiving recommendations for shopping, travel, entertainment etc. AI is using technology like machine learning (ML), natural language processing (NLP), computer vision etc., in order to carry out tasks which normally require human intelligence, for example problem solving or learning. *But how much can we trust AI for applications that are critical like healthcare, aviation or security?* To illustrate the importance of this question, imagine a computer vision system providing false information on skin cancer detection, or a self-driving car unable to detect an obstacle in front of it. Such critical applications require absolutely rapid identification, timely responses, reliability and security assurance, topics that are currently addressed by researchers.

Q. What is artificial intelligence and machine learning, terms used so often and sometimes interchangeably?

MAZ: Defining Artificial Intelligence is not easy, taking into account that the field is vast and more of a multidisciplinary entity. In a sense, AI is an umbrella discipline that covers anything related to making machines smart, i.e. a car, a TV or a software application. Along with AI we often come across another field, Machine Learning (ML), a subset of AI, referring to systems that can learn by themselves. Contrary to classic programming, which is designed to carry out specific tasks, predetermined a priori, ML algorithms are trained on a given set of data, in order to learn how to give answers to complex questions.

There are several types of ML algorithms, such as supervised and unsupervised learning, reinforcement learning, deep learning as indicative examples.

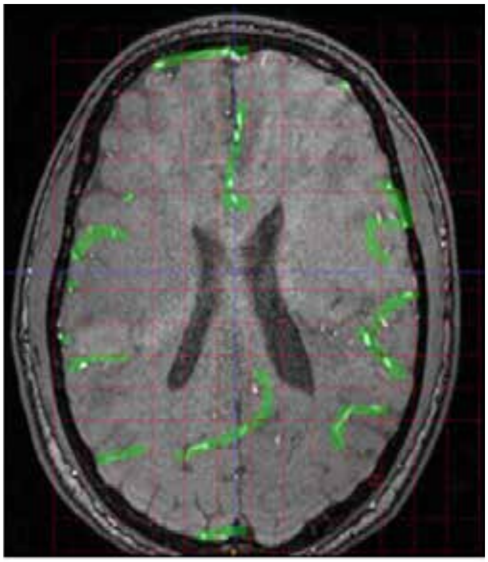
Q. What are the challenges you face in particular for AI on critical applications like healthcare and what is your approach to tackle them?

MAZ. My research focuses on the successful use of ML techniques in critical domains facing high-risk decisions, such as healthcare. Critical applications are mostly challenged by three limiting factors: data complexity, low error tolerance and reliability. I address these challenges through the development of robust ML algorithms relying on the human-in-the-loop learning principle. In other words, the goal is not to develop AI software systems that will replace physicians, but to give them tools that will help them make more informed, relevant and fairer decisions, while keeping them in the loop. One major difficulty in medicine is that data, are by definition heterogeneous, complex and difficult to use. This is why traditional ML techniques are not particularly suited. In critical applications, quality and quantity of input data are important to obtain correct results. There is a need for methods based on interactivity and on evaluating the uncertainty of the results. Especially, we are able to reduce the dimensionality of the data needed, in order for ML algorithms to be trained efficiently and still provide robust results.

This is a very useful contribution particularly for medical imaging and diagnosis, where available data for training algorithms do not exist in high quantity and require very specialised annotation processes.

Q. Could you give us examples of real-life critical applications of your research projects?

MAZ. Currently, in collaboration with UCL, we are developing interactive ML techniques, in order to ensure the timely medical diagnosis of critical situations for multiple sclerosis (MS) patients. Brain lesions detection for MS patients is a quite well known process. The challenge now is to be able to automatically detect, when such lesions touch or cross blood vessels in the brain, a situation that requires immediate medical attention. The cerebral blood vessel tree is so complex and really hard to annotate purely by using an algorithm, so humans are efficiently feeding the algorithm for high level of precision in annotations.



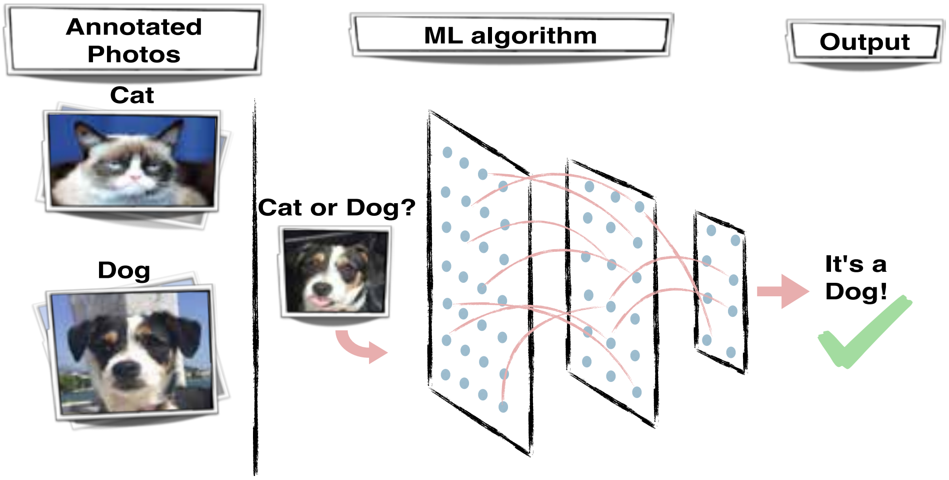
Another application we are currently pursuing, goes to the direction of personalised medicine. The idea is to use wearable devices, such as watches, t-shirts, etc. and analyse the collected physiological data of patients, in order to provide health recommendations. For example, in collaboration with clinicians in Monaco, we run a project to provide support to patients with chronic diseases, i.e. cardiac diseases, intersecting their data, with information gathered from smart city devices. Based on data for traffic, pollution etc. the algorithm will recommend to them personalised tips for their daily activities, i.e. how safe it is to go outside due to pollution, how long they should walk, whether to avoid physical strain in a humid weather etc. Furthermore, note that the methods we develop are generic and transversal, and can be relevant to all critical application domains sharing the same difficulties. For example, we are working on a project with Orange (see recent publication), to create an automated system to detect anomalies while monitoring signals on their IT system, allowing damage prevention and timely reaction to possible network issues. To reliably detect such anomalies, we use a method called “auto-encoders”, a type of neural network which is able to encode a dataset and decode it in a lower dimension, keeping the most valuable information for the reconstruction. What is considered normality in data is much easier to model and because of the larger amount of normal data available, we have a small reconstruction error. On the contrary, anomalies when decoded exhibit a significantly higher estimation error leading to their successful detection.

Q. What do you think of the future of AI and its technologies?

MAZ. There has been an exploding hype recently around AI and ML techniques and we are currently reaching a plateau for the field’s milestones. I think it is time to reflect now a bit deeper and look into what we can really achieve in the future through AI. Also, there has been a massive number of publications on the field and many unfortunate examples of mistakes found in studies. Therefore, it is highly important to increase our caution in reviewing and keep a critical mind in order to keep the essential and move forward based on more solid grounds, towards both theoretical and technical advancements on AI.

REFERENCES

https://www.eurecom.fr/~zuluaga/



Raja Appuswamy,
Professor, Data Science Department

What if there was a storage technology with the capacity to store a huge amount of data in a tiny space and essentially last forever?

DNA Data Storage

We have come a long way and data storage technology has evolved tremendously over the last five decades. Despite this, no storage technology available today is able to keep up with the rate of digital data production, or provide long-term storage of data at a low cost.

What if there was a storage technology with the capacity to store a huge amount of data in a tiny space and essentially last forever? Well, actually, there is such a medium inside every cell in your body; it's DNA! DNA is the oldest storage medium, that encodes all the necessary information for life to continue and evolve, in an incredibly tiny molecular package. Our data are our story, and in order to safely preserve it for future generations, alternative storage media like DNA might be just what we need.

EURECOM's professor Dr. Raja Appuswamy, expert in Databases and Storage Systems will guide us through, on how analog media (microform, film, archival paper) and biological media (synthetic DNA) can be used for storing digital data and what challenges we have to face, until this technology can be used in large scale.

Q. What is the actual problem for data storage?

RA. Today, magnetic tape is the gold standard when it comes to long-term data storage across all disciplines, from Hollywood movie archives to particle accelerator facilities.
It is highly denser and less expensive compared to alternatives like Hard Disks or NAND flash. However, magnetic tape medium suffers from some fundamental limitations, making it unsuitable for long-term data storage. First, it has a limited lifetime of only ten to twenty years. Thus, using tape for archival storage implies the need for data migration to deal with device failures. Second, the storage density of tape, or the amount of data you can store per square inch, is improving at a 30% rate annually.

Thus, the tape industry releases a new generation of tape every 3 to 5 years. However, tape libraries and readers retain backwards compatibility only up to two generations. Therefore, using tape for long-term storage implies constant data migration in order to deal with technology upgrades.

These migrations cost millions of dollars and put an enormous strain on archives and institutions that have to preserve data over several decades. For instance, Hollywood movie archives have already admitted that we are living in a dead period during which the productions of several independent artists will not be saved for the future due to skyrocketing data migration costs.

Q. Digital devices for storage are limited in time duration apparently. Is there no other device currently with a lifespan of a few hundred years?

RA. Of course, there are media that have a long lifespan. For instance, analog media like microfilm, microfiche, and cinema film are very durable and are expected to last up to 500 years. However, while such media have been used for preserving movies and journals across several generations, it was not clear if they could be used for archiving born-digital data. What we managed to do in EURECOM in collaboration with EUPALIA, a startup specializing in digital preservation, is to show how digital data can be efficiently stored in such visual, analog media. In fact, we map bits into barcodes that can be printed on microform, film, or even archival paper. Our preliminary solution is capable of storing 1.3GB of uncompressed data in a single 66 meters long, and 16mm thick film. Thus, analog media might be an ideal solution for enterprises with small or medium-sized archives, where a reasonable stock of a few hundred film rolls could be used to archive data for hundreds of years. Now, if a company's needs exceed several terabytes, then analog media are not the ideal solution since they are not dense enough. For massive data storage, we believe that synthetic DNA might be a good fit.

Q. It is very exciting to work on several solutions accommodating different data storage scales. Now, why DNA is ideal for massive storage and what problems does it solve that other media fail?

RA. Using DNA for digital data storage is an old idea that has been around since the 1960s. Richard Feynman even pointed out in the 60s that DNA could be used not only for storage but also as computational building block.
The main advantages of using DNA as a storage medium are the following:

1. Density: in 1 gram of DNA we could store 455 Exabytes of data. In contrast, a 3.5" hard disk drive can store 10 Terabytes and weighs 600 grams today.

2. Relevance: The density of DNA is fixed by nature, and we will always have the ability and the need to read DNA–everything from archeology to precision medicine depend on it. Thus, DNA is an immortal storage medium, which does not have the media obsolescence problem and hence, can never become outdated, unlike other storage media.

3. Durability: DNA can last several millennia as demonstrated by experiments that have the read DNA of ancient, extinct animal species from fossils that are dated back thousands of years. If we can bring back the woolly mammoth to life from its DNA, we can store data in DNA for millennia.

Given these benefits, we believe that DNA has the potential to serve an excellent archival media for storing digital data over long time periods.

Q. What is the procedure for DNA storage and what are the technical challenges EURECOM provides solutions for?

RA. Let me describe briefly the procedure of storing data in DNA molecules. DNA is a macromolecule that is made up of four types of smaller molecules called nucleotides: Adenine (A), Cytosine (C), Guanine (G), Thymine (T). So, to store digital data on DNA, we need to convert it from a binary code (sequence of bits) into a quaternary code (sequence of nucleotides). At EURECOM, we have a long history of research in developing optimal codes for classic telecommunication channels. Building on this expertise, we use concepts from coding theory to develop error-correcting codes, that can encode and decode digital data by treating DNA as a storage channel. Once the binary pairs are encoded to an A, C, G, T sequence, the DNA is physically generated using a process called "Synthesis". We are collaborating with synthetic biology companies for this purpose. We then store the synthesized DNA in DNAShell (see picture), a container developed by Imagene technologies that allows us to store DNA at room temperature for a long time period.

The reverse procedure for reading the digital data back from DNA starts with a biological procedure called "Sequencing". The synthetic DNA is sequenced using a procedure very similar to that of sequencing human DNA. The output of sequencing is a massive number (hundreds of millions) of "reads"--sequences of ACGT "strings", from which we need to decode back the original digital data. Both synthesis and sequencing are not 100% accurate. Thus, the DNA sequences that we feed as input to the synthesis step, are not the same ones we get as the output from sequencing. Some DNA sequences are overrepresented in reads with multiple copies, some have errors with spurious nucleotide being inserted, deleted, or substituted, and some DNA sequences are not present at all. To overcome these errors and recover the data, we use several algorithmic, statistical, and hardware acceleration techniques to enable fast, accurate recovery of data. Interestingly, some of the techniques that we developed for retrieving data from DNA turned out to be very much applicable solutions to problems in other domains, such as the sequence alignment step in genomics-based DNA sequencing pipelines. Thus, we are building computational biology tools to assist researchers in these other domains as well.

Q. What is the status of DNA storage now and what can we wait in the near future?

RA. Right now, the major limitation facing DNA storage is the synthesis procedure, as it is quite slow and expensive, making DNA about 100,000 times more expensive than magnetic tape. We would need radically innovative technologies for making synthesis efficient, cheaper and faster. Several synthetic biology start-ups are investigating promising, novel, synthesis techniques in a race to bridge this gap and certainly the company that solves the problem of DNA synthesis, will be at the forefront of a new revolution in computing. Being conservative, I would say that we are 10 years away from large-scale adoption of this technology. However, in the meantime, DNA storage is likely to be used in niche domains for preserving artefacts of cultural or historic significance.

Q. And a more practical question for the end! How can we guarantee that in a 1000 years a future user that would come across our data, would be able to access them? Technology will essentially be different...

RA. We are also actively working on this issue in collaboration with EUPALIA! After developing encoders and decoders for microform and macromolecules, we realised that we also need to archive the decoders in addition to data, in order to be able to guarantee successful recovery. The central idea we are working on, is to encapsulate the decoders in an emulator in such a way that the emulator itself can be stored together with the data. A user 1000 years in the future would need to write a very simple algorithm to bootstrap the emulator that would automatically execute the decoder and recover back the data.

REFERENCES

https://oligoarchive.eu/

EUPALIA: www.eupalia.com

Tired of myths vs facts? "CoronaCheck", the AI that verifies COVID-19 statistical claims

According to the World Health Organisation (WHO), along with the Coronavirus pandemic, we are also experiencing an "infodemic": a large spread of false, misleading and unreliable sources of news about the virus and its effects. EURECOM's professor Dr. Paolo Papotti, expert in data management and information quality, will guide us through how AI algorithms can be used to verify online information and especially statistical claims.

Paolo Papotti,
Professor, Data Science Department

Q. What is your approach in the fight against misinformation and how did you start working on it?

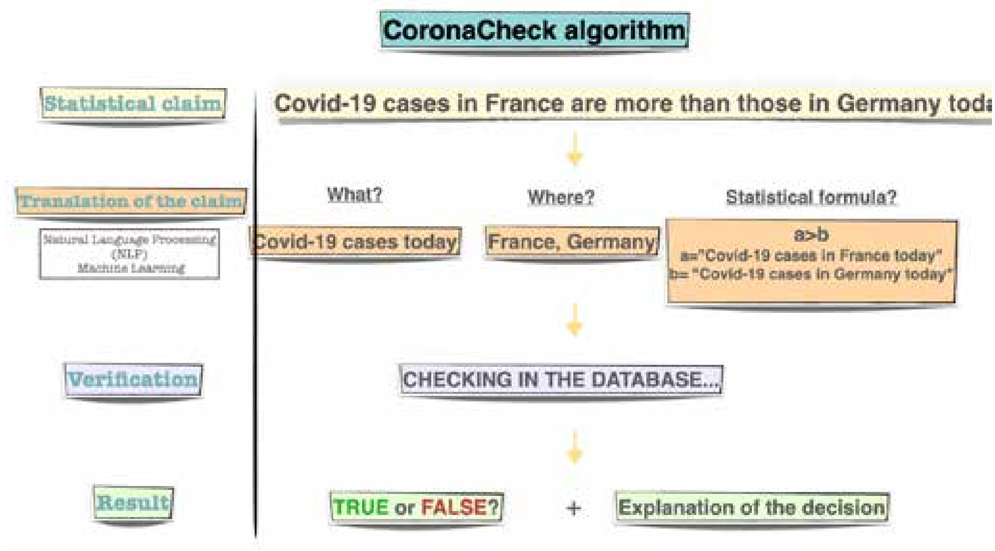
PP. I have been interested in the misinformation problem for quite some time, working on factchecking for over 3 years. The idea of factchecking is to verify if a given claim (sentence) is true or false, based on reliable sources or data as a reference frame. In this context, we started "CoronaCheck", an artificial intelligence project that aims at helping people to verify online information related to COVID-19, currently operational in four languages. More precisely, CoronaCheck checks the reliability and accuracy of the news and statistical claims circulating online regarding the coronavirus, based on the official data sources. This project is a result of the collaborations with Prof. Trummer at Cornell University and with Google, which has started supporting our research even before COVID-19. When the pandemic was in the rise and we started the lockdown, we had even more focus and motivation to work on the website and create CoronaCheck! Since then, the site has checked more than 16k claims.

Q. What types of claims we can verify with CoronaCheck and what are the hidden challenges behind the success?

PP. CoronaCheck is an AI tool that verifies the veracity of a statistical claim related to the COVID-19 pandemic. For example, "COVID-19 cases raised by 10% compared to last month in France" or "death rate in Europe is higher than in USA". Note that for a statistical claim, it is central and inherent to the statement that the information is quantifiable. However, factual claims, sentences plainly stating facts, such as "The French president has been infected with Coronavirus" are for the moment out of CoronaCheck ability of verification, but it is something we are currently working on.
Now, a major challenge in factchecking is the ambiguity of the given statements for verification. For instance, "new cases raised by 10%" is an ambiguous statement since we don't know compared to what information we should verify whether the claimed raise is true. To handle ambiguous claims, we expose several possible scenarios to the user regarding the given claim, narrowing down the most probable interpretations: You mean "news case by month or by day?" The granularity of the information plays a very important role for the efficiency of factchecking.
Finally, CoronaCheck provides an explanation for any statement found to be "True" or "False", which is what makes our tool complete and reassuring to use. It is a feature that is not obvious to implement with other "black box" Machine Learning (ML) approaches, but it comes naturally with our approach based on queries. These great results are due to the hard work of the students involved in the project, especially Mohammed Saeed, as CoronaCheck is a result of his PhD thesis.

Q. What is the principle of the factchecking AI algorithm?

PP. For our approach, which is data-driven, it is essential to create databases containing mostly trustable data, serving as the ground truth for factchecking. Now let me guide you through the factchecking process with CoronaCheck step by step.


CoronaCheck algorithm
Statistical claim
Covid-19 cases in France are more than those in Germany toda
Translation of the claim
Natural Language Processing (NLP) Machine Learning
What?
Where?
Statistical formula?
Covid-19 cases today
France, Germany
a>b
a="Covid-19 cases in France today"
b= "Covid-19 cases in Germany today"
Verification
CHECKING IN THE DATABASE...
Result
TRUE or FALSE?
+
Explanation of the decision

A person states a claim for example, “there are more cases in France than in Germany today”. The first step is to convert this claim to a SQL query in order to be interpretable by the computer. For that, we use ML techniques, namely Natural Language Processing (NLP) tools, to help understand the meaning of each claim and assign each part of the claim to variables representing the desired semantics, which can then be used by the algorithm for computations. In other words, our ML algorithm extracts each “feature” of the claim, by means of a classifier trained to link the features to a rank of possible tables, within the verified databases with a certain confidence. Of course, the mathematical formula of the statistical claim is also identified; here it would be “France cases today” > “Germany cases today”.

Having combined all the above elements together, the algorithm has narrowed down the meaning of the claim and the next step is to verify if this claim is true or false by checking against the ground truth databases. Note that verification databases are the product of hard work, collecting and integrating data on Coronavirus every day, but it is what eventually allows us to verify the requested claims. The final step is to provide the user with an answer, not only stating whether the given claim is True or False, but also with an explanation of the decision. In the beginning, the algorithm needed training since ambiguity in claims resulted in many options for interpretation.

Therefore, we generated several claims to bootstrap the algorithm, which has then gotten much better with time with more examples fed by users. We have now reached a point where our confidence of interpretation is high for many claims and our API (our algorithm exposed as a service to others) is starting to support other projects that fight misinformation.

For example, we collaborate in the Italian “Vera” chat-bot with other institutions, such as Pagella Politica, Facta, and San Raffaele Hospital, to help the public receive quick and precise answers to their COVID-19 questions. Also, we are now offering our API to infotagion.com, an independent, expert fact-checking service for COVID-19 in the UK.

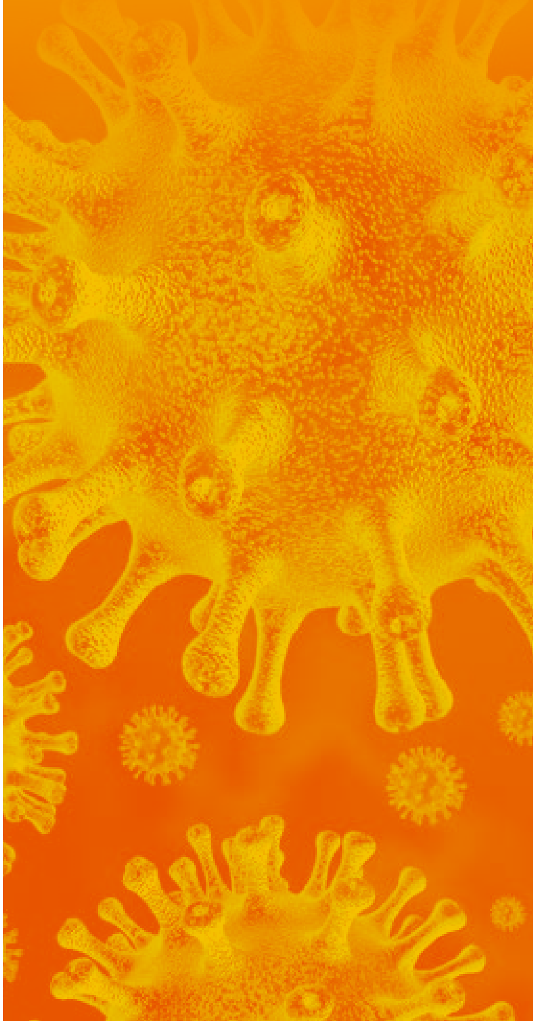
Q. Apart from COVID-19, which other field could benefit from a misinformation checking tool?

PP. Our algorithm is very generic and it suffices to be properly bootstrapped to be used for factchecking on several topics of great societal interest. For example, we have already used our algorithm, in collaboration with the International Energy Agency (IEA) in Paris, to verify claims about energy consumptions, prediction for changes in climate, and any fact that can be verified by the available data. Currently, we are writing a new project to debunk claims about EU immigration, as it is a very important societal issue used by many politicians to manipulate the public discourse. Since data is available for use to verify such claims, we would want to be able to use it!

Q. What about the future directions for factchecking tools?

PP. There is a need to support new types of claims and datasets that require solutions from an engineering point of view. For example, the next step would be to have an interface that verifies factual claims, such as “The Coronavirus is still active on surface after 8h”, since research papers have been published on this issue and they are in principle exploitable for verification. Also, we are doing an experiment to fight misinformation spread with tweets, which is a challenge on the technical aspect, because many people use abbreviations instead of full sentences. On the long term, we would like to build a unified factchecking API for various topics, allowing specialists like journalists, as well as the wider public, to verify all types of claims. Indeed, we need humans to participate in the procedure of factchecking, as we see our algorithms as auxiliary tools for more efficient and accurate performance. Google and IFCN support with grants these new lines of research.

As computer scientists, we focus right now in the task to tell whether a specific claim is true or false. This is going to give content moderators, such as in social networks, more elements to make informed decisions, in order to decide which content to filter or even remove. Very often, this a very delicate decision, since a claim that it is false may or may not be harmful. The problem of misinformation is much bigger and we play a technical role in one aspect. Governments, NGOs, media and internet companies are responsible as well, to make people more aware and able to recognise false information, through better education on this matter. It is a difficult fight that requires all actors to be actively involved.



Preserving European silk heritage through AI; historians and computer scientists are joining forces

Raphaël Troncy
Professor, Data Science Department

Although silk trade is often linked to Asian origins, it was fundamental for the European culture, linked to our history, sociology, economy and art. From furniture and clothing to tapestries and formal attires, silk was present throughout history quite often as a symbol of luxury. Following the European silk road through the centuries will lead us to a better understanding of European history, unravelling the effects of the commercial & cultural exchanges around silk to modern society.

EURECOM's Prof. Raphaël Troncy, expert in Knowledge Graphs, Semantic Web and Natural Language Understanding, will guide us through his EU research project SILKNOW, with the goal to provide insight about historic silk fabrics and AI technology applied to preserving cultural heritage.

Q. What is the idea behind the SILKNOW project?

RT. SILKNOW started three years ago, based on the idea that European silk heritage is in danger and that we would have to act in order to preserve it. There is a lot of historical information about silk fabric, mostly traded around the 15-17th century between China and Europe, making some cities in Europe quite rich, like Valencia, Lyon and some cities in Italy. Silk is a textile used at the time by high society, mainly for making clothes or furniture for different occasions. Now, historians are trying to understand the role of silk in societies of the past. This research is also expected to be used as an inspiration for today's fashion, since it is known that fashion is always reinventing itself and very often it is influenced by the past.

Another aspect of the project is to bring back to life the original techniques used for weaving silk through history. Weaving machines used cards with punch holes, in order to create the texture of the textile, while mechanically pressing with your feet a pedal of a loom. However, there is no way to know the pattern of the textile just by looking at a "punchcard"; we would have to use the card with a weaving machine. Keeping in mind that there are thousands of punchcards, we would need to activate a machine for 10 hours, just to be able to see the motif of a textile. Instead, what we propose is to construct a virtual loom that would read these punchcards and simulate in the computer what would be the drawing of the textile.

Q. Who are your partners in SILKNOW and what is everyone's contribution?

RT. We are a consortium of 9 partners representing cultural heritage, history, ICT and SME industries, all committed to keep silk heritage alive. We harvest a lot of data from plenty of museums, currently 25 museum-partners. All these museums gave us access to their silk collections, with additional valuable metadata, about production time, technique, material, location and several historical observations made by historians about the role of each silk object. Integrating various information from different museums in a common model was a challenge since descriptions were not uniformly provided, depending on the system of each museum. Now, on our side, as computer scientists, one of our contributions is to try and predict missing information on silk objects like production time and place etc., using AI algorithms to predict the missing features, based on similarities with known objects defined by historians. The common utter goal is to create a virtual museum, which would be the sum of several existing ones.

Q. What is the role of EURECOM in this project?

RT. The main contribution on our side at EURECOM, is ADASilk (Advanced Data Analysis for Silk Heritage), an exploratory search engine equipped with a spatiotemporal map in a user-friendly interface, built on top of SILKNOW's knowledge graph, literally a graph representing all the information we know. It contains around 50K textile entries with images and metadata. With ADASilk tool you can search for textiles by any keyword, but you can also search techniques by other criteria. For example, if you type "Damask", which is a type of weaving technique, you get around 20K results, with a lot of silk objects coming from different museums, that you can visualise and get more information

REFERENCES

Silk Exploratory Search Engine
https://ada.silknow.org/en

International Conference on European Silk heritage
https://weaving-europe.silknow.eu/

Follow SILKNOW on Instagram!
https://www.instagram.com/silknow.eu/?hl=en

about them. This exploratory engine, that we have developed at EURECOM, is really the integration point between a lot of technologies.

The project partner UVEG has developed a specific visualisation called spatiotemporal maps, which enables to visualise all textiles on a map, depending on where and when they have been produced. 3D maps are constructed, where every layer corresponds to a specific century and the user can see similarities and differences between silk objects. You can really see trends and evolution of textiles in terms of techniques along the time, which is a powerful tool for historians, using these rich visualisations to extend their study on silk.

Another feature integrated in ADASilk is the Virtual Loom, which was developed again by the project partner UVEG, which is acting as a digital memory of silk heritage, reconstructing the 3D image of a textile the way its weaving was done. On top of that, you are able to simulate the weaving process after changing a few parameters and finally generate a new textile, based on the old one. It is quite fascinating! Don't hesitate to test this exciting tool.

Q. Could you give us an insight on real life applications of SILKNOW results?

RT. One of the project partners is a small Polish company called "Monkeyfab", specialised in 3D printing; so, we wanted to print a dress using a 3D printer! The dress is designed by a fashion designer, using some of the textiles we have documented at our museum-partners. Monkeyfab is specialised in 3D printing of large objects, using big machines at the size of a room and the process of printing can take many days. The material used for 3D printing is a special type of plastic that looks like silk, ending up with a silk-like texture. This dress was worn by a real fashion model during a fashion show on February 27th, showing a collection by the Polish designer Patryk Wojciechowski, created with the participation of Paweł Twardo and his Monkeyfab team, using 3D printing technology!

Another aspect is that people will be able to see the evolution of textiles, virtually online on a global scale. We could make an impact on the fashion industry, if we can demonstrate the ability to print a dress for example, using this inspiration from distant past-centuries to influence the fashion of tomorrow.

Q. What are the future directions for SILKNOW?

RT. There are still outstanding challenges to overcome. For example, we are still improving how the machine can better understand domain-specific terms, since for now it is limited to encyclopaedic general background knowledge. Natural Language Processing (NLP) tools can accurately extract information from a text regarding a place or a person. But if you want to teach the machine how to extract more specific information such as the dimension of the object or the type of weaving being used, how much training data you would need to do that? In other words, we provide you with a search engine but you don't know what queries to type, since you don't know the domain specific terminology. So, we would like to have the option to search by image, instead of typing a word that you most likely don't know. Maybe you see a dress that you like or a piece of textile, so you could take a picture, upload it on the search engine and you will have more information about this. The algorithm will find results that are similar, for which we already have historical information and that we can display. SILKNOW project will finish in September 2021; now we have just completed the very first round of final evaluations and we are eventually beginning the final cycle of development until the end of the project.

Q. What is the message you want to spread through SILKNOW project?

RT. SILKNOW is a project about cultural heritage and I am part of many such projects e.g. ODEUROPA, a project which is started in 2021. Those projects are intrinsically multidisciplinary, bringing together computer scientists, historians and other various disciplines. It is true that it takes a lot of time to understand each other, since we don't share the same vocabulary, or the same timelines. For example, historians take the time to analyse things in detail and they often go more in depth. I really like these collaborations and I get very excited in understanding them and supporting them with useful tools. Our group consists of passionate individuals and also curious enough to open their minds and listen to other people's perspectives. For example, we have great collaborators in Lyon who are experts in silk and at the same time they are very open to technology aspects and wanting to understand it. Inside EURECOM, it is also important to recruit people who have this open mindset and I am very fortunate to have such people in my team, either with mixed backgrounds or very curious. For example, a PhD student who is a computer scientist, he is also in love with history, so he likes spending time discussing with historians and this is really an important part of the success.

Fashion students from EASD València, have worked together with the SILKNOW Team to propose a news fashion collection based on silk fabrics.

Punchcard used as a guide for silk weaving with a loom

3D printing of a Damask. The example shows a leaf extracted form an image of a historical textile. As it is a damask, weft and warp yarns are of the same colour

Detail of a damask reproduced with 3D model is depicted with the wireframe VL3

Digital Security Department

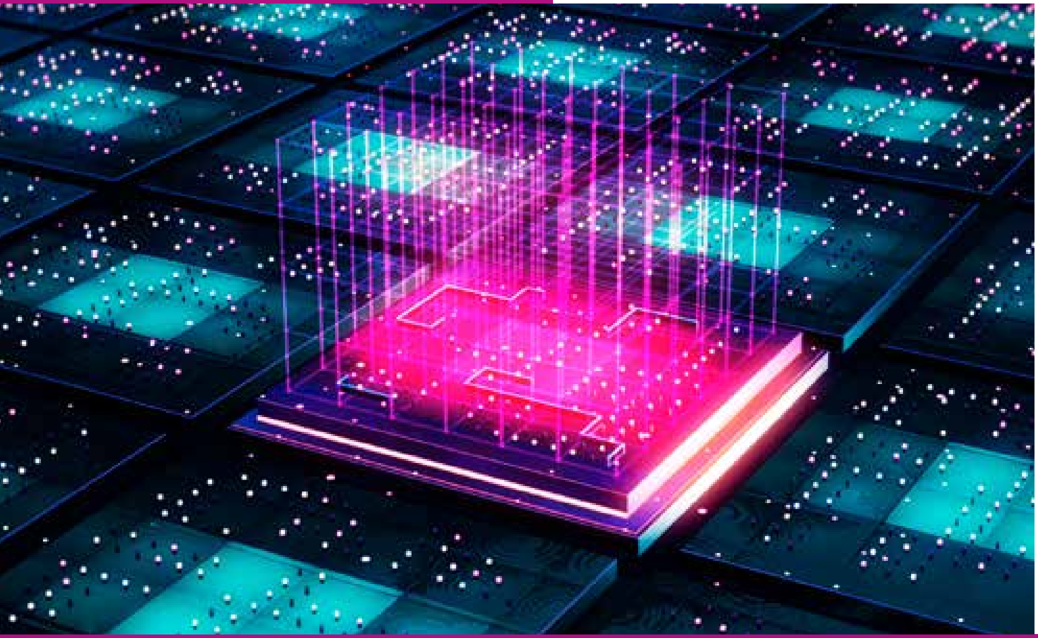
Davide Balzarotti
Professor, Head of Digital Security Department

Digital security at EURECOM: A strategic move

In 2021, Davide Balzarotti has been appointed head of the Digital Security Department at EURECOM. This new job comes with new challenges for him, but also with a new way of managing the department and the future of Digital security research at EURECOM.

Hello Davide, you started managing the Digital Security department in January 2021. Could you first describe its present activities?

We are presently working on three main thematic areas: System security, Cryptography and Privacy-preserving technologies, and Biometrics & Digital Media. There are around 40 researchers in the department: 8 professors plus many PhD students and post-docs. My job is to make the department grow in terms of topics, funding, number of researchers and students. All of this will improve its recognition in the field.



Which topics do you think the department should work on for the next 2 to 5 years? How will you guide research activities and researchers?

We have not discussed future topics yet. Of course, there are new challenges in digital security, such as supply chain security or the security of machine learning classifiers, the list of new topics never ends and keeps changing. But we also have areas that are 30 years old, and they are still fresh. And this is due to the adversarial nature of security, where solutions needs to withstand an active opponent who is trying to outsmart you.

I think the topics we will work on will depend a lot on the decisions made by the researchers of the department. In academia, freedom is important. Obviously, junior scientists like PhDs or post-docs may need some direction, but senior researchers are more independent.

Our professors are world-renowned experts in their own areas and they know best in which direction they need to go.

So, the most critical point is to hire very good researchers, especially in a small department like ours.

What is the future of Digital Security and its major challenges? This could impact research topics you might select?

Digital security is not vertical and thus it's not going in one particular direction. It is growing in every direction. Topics change all the time and good researchers can change topics easily. They know how to adapt. Every professor working in the department is recognized in his particular field. They all know how to choose the next challenges.

Experience shows that selecting the right person is more important than choosing the right topic. They are actually plenty of major topics in the field of digital security. Old ones still exist because security is a never-ending process. It is not something you can achieve with a perfect solution. Criminals always find a way to go around the solutions we design, and this in turn leads to better solutions.

Old challenges like malware, cyber-attacks on mobile phones or Web security are still hot topics. Obviously, new topics emerge: Ransomware (a way of monetizing malware), IoT, machine learning, deepfake detection etc. There are so many huge challenges to tackle that our team of researchers will have the opportunity to be selective in their choices.

For me, the best way of choosing a new strategic topic is to hire a researcher in a field that expands, but also complements, our ongoing research activities.

You've been working at EURECOM for 10 years now but leading a department is new for you. Why do you think you've been selected for this job and how do you see your new role?

It is difficult for me to say, but I guess the ERC Consolidator grant that I got three years ago helped. Plus, I have a broad understanding of security because I am some sort of generalist in the area. I have worked on many different subjects over the past 15 years, including most aspects of system security, binary and malware analysis, embedded system security, computer forensics, and Web security. But I understand the management will no longer be the same, I have to think bigger now. What you do in your own group is scientific leadership but managing a department requires a lot more: You need to be a good captain, but more than that, you need to be a good middleman. In fact, I will have to talk to all kinds of people, inside EURECOM and outside. I will obviously be close to the researchers of the department, but I will also have to align the strategy with EURECOM administration and the other departments. Plus, I'll need to find more opportunities to work with companies, find some synergies and common goals, promote some ideas. I like explaining things, giving talks, so being the spokesman for the department is something I will really enjoy.

In addition to your ERC project, you're working on big projects funded by DARPA (Defense Advanced Research Projects Agency). Will it help run the department better?

The good thing is the department has so many ongoing projects that we don't have any funding issues at the moment. Even if my schedule is already full, working on many projects will certainly help me make better strategic decisions. The ERC project I'm currently working on - Bitcrumbs, whose goal is to design tools to analyze all types of compromised systems, will give me the opportunity to work with new partners and explore more areas. I also think the DARPA CHESS (Computers and Humans Exploring Software Security) program we are part of will be inspiring: It will certainly give me valuable insights into the way funding and research work in the United States.

I also have funding from the US Air Force Research labs and a new European project on Web security with SAP. All of these projects will help me reach one of my objectives as the head of the department: **Find new partnerships and new companies to work with.** Starting a big industrial project with a large company would actually be a great achievement.

A bit like the collaboration between Qualcomm and the Communication systems department?

Exactly! The EURECOM Communications systems department did something great with OpenAirInterface. It was key to the Qualcomm collaboration.

A software platform or a similar Alliance in digital security would surely be a success for our department. It's a perfect way to transfer our research and bring different players together. Every EURECOM success is very inspiring for me and I will definitely talk more with researchers in the other departments to learn from their experience.

> Our professors are world-renowned experts in their own areas and they know best in which direction they need to go. Hiring the right researchers to tackle some of the future digital security challenges will be crucial.

PROPOLIS

Melek Önen
Professor, Digital Security Department

PROPOLIS: How can smart cities be trusted with our personal data?

The PROPOLIS project is aiming to provide solutions for privacy concerns and develop a comprehensive understanding of data privacy in the context of smart city analytics. This project is funded by a special ANR for collaborations between France and Germany.

Worldwide cities are rapidly becoming smarter by using technological advances, aiming to provide their citizens with more efficient and sustainable everyday lives. Instant information about traffic, health services, safety alerts, air pollution and community news is accessible by millions. Various devices equipped with sensors, cameras, as well as Internet of Things (IoT) are deployed across cities networks. However, an important consequence of cities digital connectivity is the vast personal data collection and processing, possibly raising individuals privacy issues. Thus, protecting citizens privacy becomes urgent and essential.

To address this critical issue, EURECOM's professors Melek Önen and Antonio Faonio, experts in Cryptography, join forces with Professor Thorsten Strufe, Karlsruhe Institute of Technology, expert in IT security. Having formed a longstanding collaboration that started at EURECOM, they are now collaborating for their project PROPOLIS (meaning in greek "For the City"), having freshly started this fall. Moreover, this consortium includes strong industrial partnerships with SAP (France) and Urban Institute (Germany). Also, the city of Antibes (France) and the city of Bad Hersfeld (Germany) are PROPOLIS sponsors. This project is aiming to provide solutions for privacy concerns and develop a comprehensive understanding of data privacy in the context of smart city analytics.

EURECOM and KIT join their complementary expertise to preserve our privacy; EURECOM with a longstanding track record of cryptographic tools and KIT with privacy preserving data analysis. PROPOLIS project has an important mission to accomplish which comes with a strong beneficial social impact. Here are the three distinct privacy preservation use cases with a direct impact to our everyday lives.

A. Traffic monitoring (use case from UI)

Traffic management is one of the hardest aspect that territories must handle. Therefore, technology is required in order to help understanding transport characteristics, such as cars distribution, speed and pedestrian circulation. To take a step further, mobile communications provided cities with enhanced means to gather more information on geolocalisation, speed, parking needs and other complex traffic patterns. However, a constant flow of citizens private data, within cities deployed networks, is being established at the same time. PROPOLIS aims to ensure the privacy of all citizens that share their data during the use of smart cities infrastructure, preventing unauthorised parties to have access to the actual private queries. “We clearly see the application of our technologies and we can realistically assess whether the privacy is protected”, says Prof. Thorsten Strufe, KIT. “Technically speaking, there is a difficult trade-off to manage between preserving data privacy and optimise performance (handling traffic congestion with minimal latency, for example)”, adds Prof. Melek Önen, EURECOM.

B. Risk prevention in public spaces (use case from SAP)

Unfortunately, in cities there are often situations where risk prevention is absolutely essential, especially in life-threatening situations. CCTV cameras in cities are operating softwares, designed to recognise a potential threat, for example an unidentified object abandoned in the street, and alert authorities. The issue that PROPOLIS project is addressing in this scenario, is the protection of the intellectual property of this software. In other words, protecting the implemented detection algorithms within city cameras, ensuring industrial confidentiality and protecting it from adversary attacks. In this context, SAP will be providing the object detection model that is protected by cryptographic tools. “Moreover, we watermark the encrypted industrial software and therefore we are able to tell whether the stolen model is ours. On the other hand, city camera footage is showing individuals faces, therefore this data need to be encrypted, such as the content is not reviewed but only the threat in question”, mentions Prof. Önen.

“This project’s use cases are driven by real life problems. For example, Antibes city, is now steal proof, meaning that if you try to steal their data, you won’t find anything in their drive. If you even try to steal the camera, you won’t be able to retrieve anything; not the data or not even the used model itself”, explains Prof. Antonio Faonio from EURECOM.

Energy consumption (Use case from UI)

Smart home applications like monitoring energy and water consumption, temperature or even critical situations detection, are being increasingly used. Identifying anomalies, for example in water or electricity usage, could help efficiently detecting a leak or any damage in the network. However, the collection of this type of data reveals sensitive information on people’s privacy, like when they go to sleep, when they are absent or what TV programs they are watching etc. “Our goal within PROPOLIS project is to detect such anomalies without interfering with privacy sensitive data”, clarifies Prof. Önen.

What are the future prospects of the project PROPOLIS?

TS: “A direct collaboration with the party that would eventually use our system, could help give us feedback on what works better, fast enough and cost efficiently. The utter goal is to serve society and enable smart cities features, without interfering with individual’s rights and by complying with the GDPR policies.”

MO and AF: “We all want smart cities, right? But we also want to protect the privacy of the individuals. Cameras are recording people’s lives, while they go shopping, where they live where they walk and we certainly don’t want this information to be accessible. With PROPOLIS project, we will allow cities to implement smart cities features without interfering with privacy preservation. What would we like to achieve? Smart and trusted cities!”

PROPOLIS will set the foundation for a strong long-term collaboration between the French and German partners within the context of the EU Cybersecurity Initiative, ensuring that the European industry can develop an edge over the competition, especially with regards to trustworthy IT.

REFERENCES

https://propolis-project.eu/

Aurélien Francillon
Professor, Digital Security Department

Welcome to avatar², the target orchestration framework with focus on dynamic analysis of embedded devices' firmware!

Towards an open source binary firmware analysis framework

Have you ever considered that most of the devices you own contain complex software, like your phone, laptop, cars, even washing and coffee machines? Software is increasingly present in everyday devices, industrial machines are more connected to the network and hardware architectures are becoming more complex, having an impact on the complexity of the software design as well. Rapidly stepping into a world of IoT, interconnected devices either for personal or industrial use, become as well more exposed to cyber threats and external attacks.

EURECOM's Prof. Aurelien Francillon, expert in software security in embedded systems, will guide us through his project AVATAR in collaboration with SIEMENS, with the goal to provide a framework for testing software in such devices in order to protect them from hacker attacks.

Q. What is the motivation behind AVATAR project?

AF. All software systems inherently have "bugs" and vulnerabilities, which can lead to many security problems, exploiting softwares and taking over the system. There are multiple ways to avoid vulnerabilities and one way to do so is testing.
Now, during the last years there is more and more software presence in embedded systems. An embedded system is a computer system, made from a combination of hardware and software, that is used to perform a specific task. They are present everywhere, in your car, coffee machine, phone and of course your very computer. For example a computer, is made of around 20 different embedded systems. Also, there are more and more industrial machines that are connected to the wireless networks.

Moreover, especially in the case of embedded systems, software becomes more complicated, written by different people, integrating different pieces of software. Also, we have a lot of critical embedded systems for driving or several industrial control systems, but also personal systems like webcams/surveillance etc. These devices are getting very exposed to the networks, which are containing an increasing amount of complicated software containing more vulnerabilities. In this context, with AVATAR project we are aiming to provide a framework for testing those embedded systems.

Q. How did AVATAR start and which are your main collaborators?

AF. AVATAR was first funded by a project we had with Google. Later, we collaborated with SIEMENS and particularly with a team that specialises in testing software for embedded systems, considering them as black boxes. This is the team that actually performs security testing for all SIEMENS devices, usually without using the documentation of the products. This process is called external are trying to find the security problems of the device like an actual hacker or external attacker would do. SIEMENS was interested in these tools to be developed, because they can be used on their real products. On our side in EURECOM, we drive all the development and the research to provide the tools that can be used on their actual devices.

Q. What is the core idea of the project?

AF. The core idea of AVATAR is simple. When it comes to security testing for software vulnerabilities, the complexity is huge, especially when A) we don't know how the hardware of the device is actually functioning, which is the case in real life very often dealing with no or limited access to documentation/source code etc. and B) handling testing for embedded systems, where we have no or little control on how the software runs. Typically, while testing a software we observe how the system is executing, try and find pieces of code that are not executed (dynamic testing), with the goal to reach the software's limits and get it to misbehave. This is a brief description on how we are able to find "bugs" or if they are exploitable, vulnerabilities. However, for embedded systems in order to efficiently test their software, we need to be able to emulate the embedded device perfectly, which is an extremely difficult task as I already mentioned. To tackle this great issue, AVATAR's goal is to create a framework in order to perform a type of an hybrid software execution, where the code is executed in an emulator but all the interactions with the real world are sent to the peripherals of the real device. For example, it is fairly easy to emulate the instructions of some code to a core of processor (e.g. ARM) but it is not possible to emulate the whole system along with its peripherals. So, in other words, we create a debugger link, connected between the PC that runs the emulator and the actual hardware.

Avatar^2 in a flash!

Avatar^2 is a python based target orchestration framework, which...

- Is capable communicating and controlling different targets
- Is designed to support a variety of targets
- Gives special emphasis on security analysis for embedded systems

This is a fairly simple idea and it is really useful for analysing embedded devices and be able to reverse engineer the software, to know what it is doing and which are the interesting parts to analyse. We are working on this for several years, having multiple contributions, and it s a topic that is is getting a lot of traction. More and more people realise this problem, and they are using our framework. We have also established a research contract with the US Army Air Force research labs, with whom we will continue working on this topic funding more research in the near future, in parallel to our collaboration with SIEMENS.

Q. Could you share the idea behind the project name AVATAR?

AF. The name AVATAR comes from the science fiction movie, where people were projecting their minds on the "Avatar" bodies on this weird planet. In the film, we have the brain on one side and the body on the other. For our project the analogy is that we execute the code program of the embedded system inside a computer constructing an emulator, which is going to execute the instructions one by one, transferring these accesses to the real devices. These pieces of software try to interact with the real world, e.g real value from a sensor, activate/open a door, by being connected to some external sensor. So in the end we don't have to emulate the real device, just transfer the the necessary information.

Q. What are the challenges to face in order to use AVATAR for real products?

AF. AVATAR is very useful when there is a security purpose or a vulnerability problem with embedded devices. However, in order to work efficiently, there are a lot of requirements. Currently, AVATAR is working well for low-end embedded devices, for example devices that are not Linux embedded, but smaller microcontrollers. We mostly target ARM processors, but there are many other architectures. We also need to have access to the firmware to be able to emulate it, as well as having access to the debugging interface. For example, very often we get devices which are not manufactured by ARM and they don't have a debug interface available. More often than not, even if all the rest of the conditions are met, many times we can't access the firmware.
The utter goal would be that companies would be able to use this software, in order to find problems in their products, being able to fix them before "bad" guys do or before even the product gets released. But we have to take into account that there are also many business and real world difficulties.

Another aspect where AVATAR could be useful is regarding the "right to repair" your own goods, a hot topic nowadays. In fact, physical goods have a lot of software in them, cars, washing machines, tractors etc. and today it is almost impossible to repair your own device, since you don't really have access to the firmware from the manufacturer. This could lead to a complete lockdown out of the device due to software issues. But even when there is no lockdown, there is a difficulty to understand how the software works and be able to modify it. For this, AVATAR could be a good way to make this easier, by automating a lot of complex tasks. There is also a movement for having security certifications for all sorts of IoT and physical goods, especially in the US. Tools like AVATAR could take part in this context and help these movements.

Q. What are the future directions of AVATAR?

AF. In the future, we are going to continue working towards improving the system, mostly including more complicated Linux devices as well. We also want to improve the scope of AVATAR and make it more used, find more use cases and promote it to industrial parties, that are going to be interested in such approaches. For example, the US Air Force is leveraging avatar^2 in their research and testing environments. That's why they found us and they want to support the development of AVATAR, which is an open source tool.

To conclude, the goal of avatar^2 is to be a versatile platform which can be used for many purposes on many different devices. We want to integrate various useful tools to provide a platform for testing embedded devices.

Researchers are working towards **privacy preservation!**

Nick Evans
Professor, Digital Security Department

Researchers are working towards privacy preservation! Voice interactive systems are omnipresent these days, since voice is one of the most natural means to interact with automated systems. More and more we have devices at home with smart speakers, always listening to our conversations and probably recording privacy sensitive information, that could be linked back to you. It is now essential more than ever to consider "voice anonymization", which means removing the identity and suppressing all the privacy sensitive aspects in a speech signal.

EURECOM's Prof. Nick Evans, expert in speaker recognition, privacy preservation and biometrics will guide us through the VoicePrivacy Challenge Initiative, a project dedicated to anonymisation and identity protection in voice speech signals.

Q. What is the context behind VoicePrivacy challenge?

NE. The effort within the community to look at privacy preservation within the context of speech and language technologies is quite recent. There is a history of work but the topic has not really taken off until recently, meaning the last year. Realising this fact and given also the nature of speech and at what degree it can contain and reveal privacy sensitive information, this was something our community had to address.

So, we have relatively recently started exploring the privacy aspect and in this context we launched the VoicePrivacy Challenge Initiative. It is an international collaboration mostly between French and Japanese partners and it is also supported by an ANR project called HARPOCRATES (ANR-19-DATA-0008). The participants in the challenge had to use standard database, protocols and metrics to benchmark different solutions for anonymisation.

These types of benchmarks and challenges are really the drivers that spearhead progress because without standard protocols and metrics, it is impossible to compare different solutions. Each lab is going to evaluate their solution on a different database/protocol/metric and then we won't know whether any claimed improvements to performance come from the algorithm or simply because they are evaluating their algorithm on a different database. So, it is a community led initiative that takes the form of a challenge, where different labs around the world developed solutions to anonymisation and then they submit to us their speech data that is processed using their algorithm. In the end, we evaluate the anonymisation performance using our metrics.

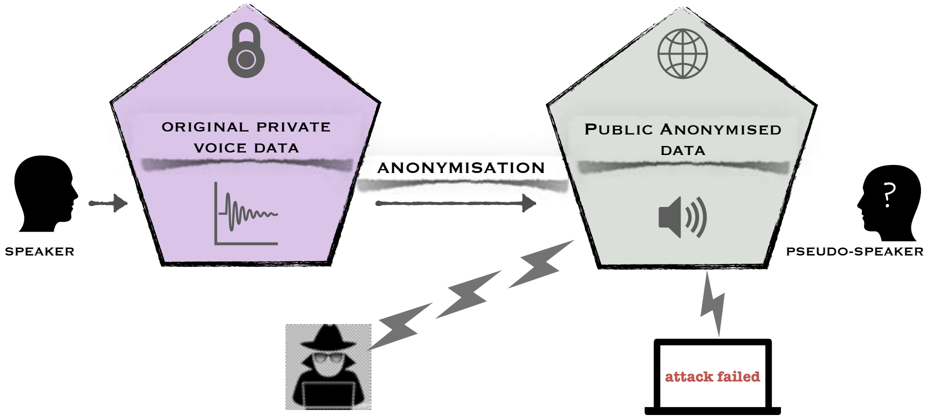
Q. What is the role of EURECOM in VoicePrivacy Challenge?

NE. My lab is the Audio Security and Privacy group, in the Digital Security department at EURECOM. In our group we are mainly looking at anonymisation through two angles. The first one, is to try and stop an eavesdropper intercepting your speech data. The second one, is to try and suppress any privacy sensitive information from your speech data, without revealing any sensitive information and without any encryption needed. In other words, we are working on voice anonymisation, the process of suppressing the identity from the recording of a speech signal, removing information for example of your gender, your age, your job titled even your emotions while speaking. An obvious requirement with an anonymization process is that we want your voice to still be either intelligible, meaning that a human or a machine can understand what you are saying or natural, meaning that a human being would listen to what you are saying and he wouldn't know if your voice has been treated by an anonymisation system.

On the other hand it would mean that a speech recognition system, trying to understand what you are saying would still perform reliably and still interpret your message or your request even though it's been treated to remove your identity.

Q. How can we define "identity" in a speech signal?

NE. Speech signals are quite rich in terms of the different sources of information that they contain and identity is one of them. These sources of information could reveal for example your gender, age, quality of voice,



emotions etc. Voice is a biometric and it can be used in order to determine who is speaking. Humans recognise the quality of voice of people they know quite well and machines are able to do the same thing and learn to recognise your voice, in the same spirit that machines can recognise your face or your fingerprints.

But let me explain about biometrics and identity. For example, with DNA the chances of confusing two people are one in a billion. Now, face, iris, fingerprints, voice should be also very distinguishable. We can use these characteristics, called biometrics, to differentiate one person from another and be relatively confident in the result. However, biometric characteristics are not unique. In fact, each biometrical characteristic has different degrees of distinguishability. There is often a misleading narrative in the media on being able to uniquely identify a person through such biometrics and we have to keep this in mind.

Q. How does anonymisation work?

NE. Regarding anonymisation, in terms of the VoicePrivacy Challenge, we want to develop an algorithm that would disentangle the identity information from all the other sources of information within the speech signal. Then, the goal would be to resynthesize the speech signal with all those other components except the identity. There is an old argument claiming that if we are able to suppress the identity from a speech signal there would be no privacy concerns. Here, I would like to clarify some usually misused terminology. We often talk about identification, authentication, recognition and verification and we have to keep in mind that they all have different meanings. The most general term is recognition and then we have identification on one hand and on the other hand authentication or verification.

Identification means that you would present your biometric to the system and it has to determine who you are.

Authentication/verification means that you would present your biometric to the system and it replies with a yes or no.

In the case of identification there is no claim of identity, so we have "1 to N" comparison. In the case of authentication/verification there is an inherent claim of identity and so an "1 to 1" comparison. More generally, we haven't got this far into looking at specific applications for voice privacy, but we are using authentication/verification type of experiments to determine how successful the anonymisation system is. More specifically, we run some experiments with a database of natural voices and we would judge the performance of a speaker verification on that database. Then, we would anonymise all the voices and then we would run the same experiment. Normally, the performance of the speaker verification system would deteriorate. If it deteriorates to the point that the error is around 50%, it means that the system is guessing and it can not verify with any confidence your claimed identity.

Regarding the tools we use, there is a variety of different solutions ranging from very traditional signal processing tools, to very complex deep learning solutions that are based on neural acoustics and waveform models. Both methods can distort the spectral envelope of your speech signal, such that speaker recognition systems don't work anymore. Most people in the community are working on deep learning approaches, but I personally believe that we can also do pretty well without them.

Q. What would be an example of real life applications for such voice anonymisation systems?

NE. Imagine you attend a party one day with friends and maybe you drink a little bit more than you should. Your activities are recorded and posted online on social media. Now imagine your employer gets hold of this and causes you problems. Now, if that data was treated with anonymisation, your face will be masked and your voice treated such that we could not tell if it was you. In other words, if you want to publish some data online, but you want to protect your identity, anonymisation is a solution. So, publishing content on social media would be one example of anonymisation. Let's say now you are using automated voice interactive systems when you call your bank or your insurance company. It is possible that anybody who can get a recording of your voice could derive privacy sensitive information from that recording. The point is that with GDPR, there is a drive that is putting control of privacy into the hands of the individual. So, you should be able to decide for yourself, what information is being used by companies and for what purpose. But if you provide the information without protection, you are trusting tech companies not to misuse your private data.

Q. What are the challenges for succeeding in anonymisation?

NE. One of the big challenges is how to actually measure privacy. In fact there is no broadly accepted legal definition of privacy. One can wonder how the GDPR came to force when no one really knows what privacy means.

One other topic is computational complexity, keeping in mind that many of the systems must operate in real time. You might have an app installed on your phone that could anonymise your voice when you call your health insurance provider or your tax office. Many state of the art solutions can not run on a mobile phone, since there is too much processing especially for the deep learning type techniques which are just too heavy on computation. Whereas other lighter computational methods might not provide enough privacy preservation but you might be able to get them to run on a mobile device. So how do we strike a nice balance between utility on personal consumer devices and the level of privacy preservation that is provided? That is a future challenge!

Q. What are the future directions of this field and your message as an expert?

NE. As mentioned before, understanding what privacy means into the context of speech signals is very important. Andreas Nautsch, a research fellow at EURECOM is an expert on speech data within the legal community and misconceptions about legal implications within the speech community, having written a very nice article on this topic. How anonymisation solutions should be integrated within existing speech technologies is an open question. How existing speech technologies should be adapted given that we may need to include anonymisation. Because if we anonymise speech, it is possible that tasks like speech recognition may not work so well anymore. We also have a lot of work to do in assessment to compare different solutions reliably. A central goal of the VoicePrivacy initiative is to bring answers to these questions. The goal is really to raise awareness that privacy is important within speech technology and that we can address the requirements of GDPR. But it is important to realise that we can't put all speech technologies into one bucket. Different speech applications would require different privacy preservation solutions. Anonymisation is not the silver bullet; the field is much broader.

REFERENCES

https://www.voiceprivacychallenge.org/
Introducing VoicePrivacy Initiative
https://arxiv.org/pdf/2005.01387.pdf
Article on Privacy Definition
https://www.isca-speech.org/archive/Interspeech_2019/pdfs/2647.pdf

New Talent

Massimiliano **Todisco**

Professor, Digital security Department

Q. What is your academic trajectory?

MT. Originally, I am a physicist and I have a PhD in music signal processing. My research interest spreads across many fields like signal processing, statistical pattern recognition and Machine Learning. After my PhD I stayed in the University of Rome for another 3 years as a postdoctoral researcher. I am also a professional musician with a degree in trumpet from the Conservatory of Rome, so this is why I am so interested in sound and music engineering.

After having worked with music signals, I started moving towards speech signals in my later years in Rome. We started writing a proposal for an H2020 project which was successful, where Prof. Nick Evans from EURECOM was involved as a speaker recognition expert. After this in 2015, I had an offer from EURECOM to work with N. Evans as a postdoc for 5 years working on speaker verification. In 2020, I got a position of assistant professor here at EURECOM.

Q. What are the expertise you bring to the Digital Security department?

MT. Music and speech communities, even though they work with audio signals, don't communicate much between them. So eventually, I brought my expertise and tools that we used in the music community and applied it for speech signal.

In fact, I propose to EURECOM a new kind of research. In the Digital Security department, I proposed to combine cryptography and biometrics together to find solutions to both. Namely, apply cryptography algorithms to biometric systems for the recognition of people. We already co-supervise 2 PhD students in combined topics between crypto security and biometrics. However, from my signal processing point of view, I can bring expertise to crypto algorithms. Biometric recognition is based on Machine Learning algorithms, which are not traditionally designed with privacy and security issues into consideration. Legislation moves forward and there is a requirement regarding security that we have to consider. Privacy preservation techniques are urgent to change adapting to the way the biometrical signals are processed. The actual challenge? When we apply cryptography on Machine Learning, the accuracy drops a lot. So, we need to adjust the cryptography tools also, so that they can be combined with ML.

Q. What are the highlights of your research here at EURECOM?

MT. Around 2016, we were able to develop here at EURECOM the state of the art framework features and structure for anti-spoofing. Today, it is very well known and widely used by the community. For this achievement, we got a best paper award in one of the top conferences, Odyssey 2016, in Bilbao. **Also, this work got the best paper in the Computer Speech and Language journal during 2015-2019. I am very proud of these results!**

Q. What are you current research projects that you are working on?

MT. The H2020 RESPECT project on privacy preservation, multi-biometric personal verification and anti-spoofing, is currently running, where I am the coordinator. The objectives of the RESPECT project are to simultaneously ensure both security protection and privacy preservation. At the same time, recognition accuracy and reliability will be

Daniele **Antonioli**

Professor, Digital security Department

Q. What is your academic trajectory?

DA. My name is Daniele Antonioli, I'm from Italy, and I've joined EURECOM as an Assistant Professor in June 2021. I obtained a PhD in Computer Science from the Singapore University of Technology and Design (SUTD). During my doctorate, I visited the University of Oxford and the Helmholtz Center for Information Security (CISPA). Before joining EURECOM, I worked at the École Polytechnique Fédérale de Lausanne (EPFL) as a postdoctoral researcher.

improved, thereby boosting confidence in future multi-biometric audio-visual recognition technologies and realising their societal and economic benefits.

We are also running a Marie Curie Early Training Network (ETN), TReSPAsS. TReSPAsS-ETN is a consortium of seven universities, supported by seven industrial entities, located in France, Germany, Netherlands, Switzerland, Spain and Belgium. The network's goal is to train Early-Stage-Researchers at the intersection of biometrics, attack detection, privacy and security, and legal and ethical issues. On the research side, TReSPAsS-ETN project will deliver a new type of security protection, through generalised presentation attack detection technologies and privacy preservation, through computationally feasible encryption solutions. **It is important also that in this project, we don't only deal with the engineering aspects but also legal and ethical aspects of security.**

Biometrics, attack detection, privacy and security, legal and ethical problems are rising. The most difficult problem we have to solve is to integrate cryptography algorithms inside biometric systems for recognition of people. The aim is to find computationally feasible approaches through deep learning and cryptography for using biometric data for personal recognition.

Q. What are the real life applications for your research?

MT. Currently, I work a lot with voice but also other biometrics. For example, in order to do a bank account transaction by voice using a mobile phone, we need high security. The system has to recognise the user with high accuracy and then you can proceed to the transaction. In fact, our results can be used to all the applications that involve the identification of people.

For example, home voice control systems like Alexa. In order to buy stuff online, it is important that Alexa or Google recognise you as the actual owner of the account. In these applications, there is always a threat of an attacker trying to impersonate you. Not only regarding security but also privacy, since people can know what you buy, so your voice has to be protected.

Q. Do you have common research projects with other departments in EURECOM?

MT. Yes, I also work with Maria Zuluaga form the Data Science Department, on Covid-19 detection, through voice and cough. Maria's group has expertise on AI applied on health applications. This project started in 2020 and we are also in collaboration with a company in Spain. Our partners in Spain, are collecting data from patients with Covid-19 in hospitals, having confirmed they are positive with a PCR test. We would like to achieve Covid-19 detection using a mobile phone! We will also participate in a challenge in Interspeech.

Q. What is the take home message you would like to give?

MT. Now, I am working on biometric recognition and privacy preservation. I think that the main problem now is that people do not trust these systems yet. We need to help people with the idea that these systems are secure and don't steal your identity or your thoughts. On the contrary, the tools we develop are here to hide all the personal sensitive information from attackers, like gender, age, etc.

REFERENCES

https://www.trespass-etn.eu/

http://www.respect-project.eu/

https://www.asvspoof.org/

https://www.massimilianotodisco.eu/

Q. What is the expertise you bring to the Security department?

DA. My research revolves around two areas: security of information technology (IT) systems (e.g., IoT, wireless networks, embedded, and mobile) and security of operational technology (OT) systems (e.g., industrial control systems and cars).

My research complements and extends the research interests of Prof. Balzarotti and Prof. Francillon in several key areas. Examples include wireless systems, embedded systems, protocol analyses, applied cryptography, and industrial control system security.

Currently, I collaborate with professors from EURECOM's digital security and wireless communication departments and local and international companies doing system security research. Being very curious, I'm sure that many more collaborations will arise over time.

Q. What made you choose to come to EURECOM?

DA. I choose EURECOM because it provides a nice mix of work-life conveniences. As a scientist, I've joined excellent and internationally-known researchers in relevant areas, including system security, telecommunications, and data science.

As a teacher, I have the privilege of interacting with talented students from different cultural and technical backgrounds.

Regarding lifestyle, EURECOM's location is excellent and enables one to enjoy the French riviera, inland, and alps.

Q. What are your future goals and your message as a new Assistant Professor at EURECOM?

DA. We have to (quickly) address new and complex system security challenges resulting from the mix of information technology (IT) and operational technology (OT) systems. Our main goal is to develop original and valuable solutions to tackle such challenges to secure current and next-generation pervasive technologies.

REFERENCES

https://francozappa.github.io/.

ANTONIO **FAONIO**

Professor, Digital security Department

I did my PhD at Sapienza University of Rome in Cryptography, and then, my first post doc was in the Cryptography group of Ivan Damgaard in Denmark, working on a branch of Cryptography called Leakage and Tamper-Resilient Cryptography. For my second postdoc, I moved to Spain at IMDEA Software Institute in Madrid, and worked with Dario Fiore, where I shifted my research focus, mostly to applications for cryptographic protocols. In July 2020, I arrived at EURECOM as an Assistant Professor in Cryptography, in the Digital Security Department!

Ensuring the safety of our digital lives with new cryptographic models

Q. What is the expertise you bring to the Digital Security department?

AF. At EURECOM, there is already a very strong expertise in Multi-Party Computation (MPC), for many ML applications which are actually used. One of my expertises is on Leakage and Tamper-Resilient (LTR) Cryptography, a field that studies the provable security of cryptographic schemes when the attacker could exploit the physical properties of the devices where the protocols are implemented. For example, in a timing attack setting, in which the adversary could compromise the system by measuring the time needed to perform a cryptographic algorithm, e.g. return back a signature to infer the secret key. With the LTR approach, we define more general models, and prove that schemes are still secure, even though attacks could be stronger.

In the last years, I have also been working on efficient zero-knowledge proofs, which are fascinating cryptographic constructions, by which one party can prove the veracity of a statement, without conveying any further information.

Q. How could we use such security schemes in real life?

AF. There is an example drawn out of most people's lives. So, imagine you want to rent a house and the owner of the house asks: do you earn enough to pay the rent? Usually, the rule is that you must earn at least 3 times more than the rent to be an eligible candidate.

What is requested afterwards by the owner is a proof of income; so we provide him with our payslips in order to validate our salary. But in fact, we reveal much more. The owner gets to learn the exact amount of our salary and not just if the 3-times bigger condition holds. Now, if the owner is a malicious person, he could raise your rent for next year if your salary is much higher.

Zero Knowledge-proof could be used by the renter to show that you earn more or equal to what is required for rental without needing to disclose more information. At the end of this protocol the owner of the house is covered and we don't reveal our privacy data to a potentially malicious person.

Q. What are your future goals and what message you would like to send as an expert in cryptography?

AF. For the next 3-4 years, I would like to start investigating questions about MPC protocols with reasonable security guarantees even when all parties are corrupted. I think that we are really doing better as a society and especially in the last few years, when it concerns our privacy online. We have a choice for each website to accept cookies, people refuse to use applications that are not end-to-end encrypted, and they are more cautious about what we post on social media. Cryptography could help us keep on having a nice way to communicate, and to collaborate in many different ways, with people online but making sure it is secure and preserving the communication as intended to be shared.

PUSHING THE FRONTIER OF 6G NETWORK

"Prof. Derya Malak brings a very original angle to 6G wireless research. Her idea is to study the fundamental limits of distributing computing which wireless networks enable when a lot of wirelessly connected devices with computing capabilities can pool their resources together. It is a promising avenue towards green (low energy consuming) computing "

DERYA **MALAK**
Professor, Communication Systems Department

Q. Could you describe briefly your academic trajectory so far?

DM. Originally from Turkey, I did my bachelor's at the Middle East Technical University, Ankara, and my master's at Koc University, Istanbul both in electrical engineering. As a student, inspired by my professors I decided to follow an academic career. Hence, I pursued a Ph.D. in wireless communications, specifically on device-to-device communications, at the University of Texas, Austin. At that time, I realised that I would like to diversify my background, leading to my postdoc on information theory concepts at MIT with Dr. Muriel Medard, a great role model for me. In 2019, after my postdoc, I got a faculty position at Rensselaer Polytechnic Institute (RPI) at New York state, where I worked at the intersection of information theory and computation.

Q. What is the expertise you bring to the ComSys department?

DM. My expertise is in the area of communications, and I am intrigued by the use of communication networks towards next generation computation, meaning the use of communication networks for the purpose of computing and tasks. Usually, we use networks with specific objectives, using the available resources effectively, with minimum information transmission, in order to recover what we need. In the future, I am hoping to form collaborations within the ComSys department, at the intersection of theory and practical implementation.

Q. What are your future goals and your message as a new Professor at EURECOM?

DM. I think that the solutions to the problems I am looking into lie within the intersection of already well-established fields like information theory and graph theory. In 10 years, I would like to contribute to the understanding of fundamental aspects of the computational limitations of networks. Even if it would be a small contribution, it could help towards building a field of computation in networks, as Shannon did with communication in networks. A researcher's life is a choice that requires a lot of dedication and perseverance; in this job, we have to fail a lot - 10 times (if not more) to be successful once! We can change the world starting with things within our reach. For example, inspire students, and touch their lives to pursue their research interests guided by our expertise. If you are passionate, you should not give up early on the path of academia; given enough time, you will eventually get where you want to be.

A new era for 5G and future generation networks

Raymond Knopp,
Professor, Head of Communication Systems Department

In January 2022, Raymond Knopp has been appointed head of the Communication Systems Department. His vision is to bring additional expertise to the department, by recruiting new faculty members, that will help shape new research directions. Historically, there are two basic poles in the department; communication theory and network systems. Recently, Professor Knopp has started a third one in standardization.

Q. Could you briefly describe the Communication Systems Department at EURECOM?

RK: The Communication Systems Department is one of the oldest at EURECOM and has generally maintained the same spirit of mixing theory and practice since the beginning. It is a place, where we always try to combine fundamental studies in communication theory with prototyping of schemes inspired by theoretical results. Overtime, the department was restructured into two groups; one concentrated in communication theory from for physical layer communications to network modelling and performance analysis. The other focuses on systems research and prototyping also covering multiple protocol layers. So, concretely the spirit of the department is to balance those two components and propose holistic approaches to future systems.

Q. What are the most important projects and applications?

RK: First of all, the ERC grants that the department has had over the last 5-6 years represent the flagship projects of the department. It started with David Gesbert's ERC Advanced grant, then Petros Elia's ERC Consolidator grant and more recently Marios Kountouris Consolidator grant. Those grants put us on a path towards more disruptive areas in the field. In the coming years we will try to maintain this level of high-impact projects in communication theory. Regarding the network systems group, the key projects are the ones that integrate us with the international research community in systems research, like EURECOM's partnerships in large EU projects consortiums and the development of OpenAirInterface. Projects like 5G EVE, which finished last year, have really established EURECOM as one of the key experimental wireless sites in Europe, interconnected with other similar sites. Follow-up projects building on that, like 5G!Drones, 5G-Victori, 5G-Records and other 5G-PPP projects, where they were using EURECOM's site to integrate with partners that are develop key applications on top of 5G technology.

Q. ComSys department has a particular high rate of ERC grants and that is very impressive. What do you think is the reason of this success?

RK: First of all, the department is very recognized for its contributions to communication theory over the years. Since the very beginning of this department, there were hirings of a few very high level professors in communication theory. For example, Prof. Dirk Slock has been with EURECOM since almost its foundation, and he is an extremely recognized world-class researcher in signal processing and helped shape the department's reputation of excellence. So, I would say that the success comes from the original vision of the founders of EURECOM and also from the fact that the department was able to put together all the right ingredients, not only regarding the high excellence in research but also the very relevant research topics for industry. The ERC project Perfume, that David Gesbert put together 6 years ago, was very solid from a theoretical perspective but it also addressed very important problems, leading to what is called 6G today. Similarly, younger researchers looked at theoretical questions that also addressed important practical problems. That is one of the reasons that we have such a high success rate.

Q. What is the future perspective and research directions for the department?

RK. We are not going to change very much, since everything is in place. We just have to follow the strategy we have been doing for years already, keeping the two poles of communication theory and network systems stimulating each other. We might try and diversify some of our funding methods a bit more. We have many industry contracts one on one with key industry players which could increase as we enter the 6G era. By diversifying the funding sources we protect the continuity of our research in low public funding periods since European and National projects always come in waves. At the beginning of the department, we heavily targeted European projects and at some points we suffered because of these funding waves. The other goal is to maintain or even increase the ERC grant success rate in the future. We will also try to increase the collaborations between the departments at EURECOM. As a first step, we have started mutualizing our computing infrastructure, encouraging new collaborations and inter-departmental projects. We have a few examples already happening with the Digital Security department and we should keep going towards that direction, which was also put forth by David as the new director of EURECOM.

One of the challenges that we have is to motivate younger students for communications theory and networks. So, our goal is to find ways through teaching, especially now with EURECOM's new engineering diploma, to overcome the difficulties in the field, due to highly technical requirements. We would like to increase the number of PhD students, since history shows that most EURECOM's young doctors don't go on in academic careers, but they land high-end industrial positions in technology development or standardisation.

Q. What are your long term goals as Head of the Department of Communication Systems?

RK. We have some faculty positions to be filled in 2022. So, this is giving us an opportunity to shape a new direction in the Department's research. There will be new fields that we will be looking at, especially communications for cyber physical systems in the network systems group. But also, there will be new professors in network and communication theory either in a classical or a more futuristic area. The other aspect to look at, is the matter of skills. Today we are a heavy electrical engineering profile in the department. I would like to have one very strong computer science faculty member, in order to bring a different vision.

Also, since last year, we are part of a large European research infrastructure called SLICES, which is a pure academic instrument. We will be interconnected with other first rate EU labs, similar to those at EURECOM in Europe, to harmonise our ways of doing experimental research in communication systems and in computing. Although we are only in the preparatory phase, this structure is planned to provide research and training services over 20 years and will bring EURECOM closer to other labs in France e.g. Inria, CNRS, IMT but also internationally.

As I said before, we already have the two basic poles of the department, communication theory and network systems and we recently started a third one with the standards group which is the result of the Patent Factory project with Qualcomm, France Brevets and IMT. The standards driven research group should now feed from the research results coming from the two basic groups of the department and push some key ideas into 3GPP standardization, as a means to find more funding for basic research through monetization of patents. For the moment this group is limited to our department but ideally, I would like for it to grow more and other departments could benefit as well.

David Gesbert,
Prof. Director of EURECOM

How could the future wireless networks look like?

Nowadays, wireless networks are mainly configured in a centralised manner, where the whole traffic is managed from a central point. In fact, a central unit collects and processes all the information, in order to make decisions on resource allocation, a process often introducing extra delays.

However, in the future we can imagine having billion devices from phones and tablets to industrial machines and drones, struggling to use the spectrum, causing a lot of interference, but also demanding instantaneous decisions, especially in the case of critical applications. Naturally, we can wonder, could we exploit part of the capabilities of such devices at the edge of the network by deploying a network in a distributed way, instead of relying on a centralised one?

Prof. David Gesbert, holder of the ERC Advanced grant PERFUME on Smart Device Communication, will guide us through the main outcome of PERFUME project: an emerging distributed network design, featuring intelligence and decision making at the edge nodes.

Q. What is the context and motivation of your finished 5-year ERC Advanced project, PERFUME?

DG. Our goal was to design how future wireless networks will operate, going towards the future 6G (beyond5G) technology, making use of all the capabilities of the network closer to the end user. In other words, exploit the memory, the computing power and transmission capabilities of phones, tablets, PCs but also robots, which are all equipment residing at the edge of the network. Currently, the terminal equipment is viewed as passive devices, running applications in a way that does not require local intelligence, underutilising their potential capabilities. Also, in the future the rapid increase of automated industrial manufacturing, using robotic devices to perform essential tasks, will rapidly increase the need for timely decisions locally, which are extremely hard to achieve in a centralised deployment. Indeed, those devices have computational power used currently to run their applications, but not yet for the benefit of the network. So the idea is, why don't we use the network in a distributed fashion and exploit part of the devices capabilities for the service of the network, instead of relying on a centralised network, where we need a unit that controls all traffic?

Q. What is the difference between a centralised and a distributed network?

DG. In a centralised network, the central point would have indeed all the information needed to make a good informed decision on how the resources would be allocated to the devices of the network (power transmission, frequency, bandwidth), but it will take much time to collect it. So, by the time the decision is sent back to the user, it could be already outdated, since networks are not static. For example, imagine you are in your car, moving around, with an environment changing constantly, therefore if you take too much time to make a decision, there would be too much latency. On the other hand, in a distributed network the decisions could be very quick, since there is no need to centralise everything. However, each terminal only has access to its local neighbourhood information, "seeing" maybe a few other devices, without knowing exactly what the other devices are going to do. So, they are going to try and cooperate based on partial information. If I have to give an analogy, imagine you are playing football with a team and you need to score a goal, but suppose that the players are playing blind folded and they don't actually see the ball. So they have to operate with hearing and feeling but not seeing.

REFERENCES

http://www.ercperfume.org

PERFUME Publications
http://www.ercperfume.org/pubs/

Now, they have to progress to the opposite side of the field, pass the ball to each other and score the goal. Well, it is not easy to cooperate when you don't know what your teammates around are doing! The same problem applies here. Devices would like to cooperate with each other to minimise interference, but they have limited communication capability between them, needing a robust coordination.

Q. Could a distributed network configuration improve the network performance then?

DG. In fact, between the two network configurations we end up with a trade off: more timely decisions but based on partial information (distributed) versus more complete information but decisions with more delay (centralised). In other words, either we centralise all the information for all the devices, push them to the cloud to make a decision and send it back to the network, or we maintain the decisions at the local level. For a critical mission, needing an immediate response as accurate as possible, can we know which is the best design between the two?
Well, there is no easy answer! For that, there is a need for extensive analysis, but we were able to show that distributed systems would do just as well and even better than the centralised ones. Then, the choice depends on the exact application. The key idea of the PERFUME project is to develop decision making algorithms for connected devices, made to be robust with respect to the fact that the information is noisy, local and not complete. We developed a new theoretical framework for this problematic, defined the fundamental limits, designed machine learning (ML) algorithms and implemented them into simulation software.

Q. So, what is the main outcome of PERFUME project?

DG. Apart from the design of ML algorithms for efficient and robust distributed networks, there was a part in my proposal, which at the time I did not think that it would be central. But things turned out differently as they usually do in research. We experimented with the above described algorithms that we designed, using terminals that are not phones, but in fact robots, like drones. Imagine that we have drones that carry radio equipment, which would essentially make them flying base stations, creating a flying network. Then, we would need that drones to position themselves and fly at the appropriate positions, so that they can serve the communication needs in the best possible way.
Today when we make a phone call, we receive it through the nearest cell tower coming from our operator, but it is a fixed one. Now potentially, the drones would be able to move around freely, and put themselves always at the optimal position, so people can have the best communication quality. Drones can collect constant measurements and then our algorithms convert those measurements into an optimal path. Essentially the drones will be flying autonomously to the optimum location in the sky, such as we could obtain the best possible coverage at any time. We came up with a lot of interesting problems when using radio equipments and so we formed a lab of 10 people and this ended up being a large part of the project. The application to robots drew a lot of positive attention and this work became central, with half of the people on the team working on drones.

Taken together, I would say we were able to provide two main outcomes:

1. ML algorithms and conceptual level on how to have efficient and robust distributed systems at the edge of the network.

2. Flying radios, through connected drones programmed to optimise network coverage.

Q. Drones are a very interesting application of connected devices at the edge of the network. What are the challenges for actually deploying flying connected drones?

DG. This is new in a sense that there is no product like that, because indeed there are a lot of challenges. First of all, it is not easy for people to accept the idea of having drones in the sky, finding it threatening and untrustworthy. So, the first obstacle to overcome is social acceptance.
However there are some applications for which flying drones would be very useful, such as rescue missions and missing people localisation. Another one is to provide extended coverage in case of people gathering for a sporting event or a concert and there is not enough bandwidth for the service providers to give internet access to everybody there. We could very quickly provide extended coverage capacity by flying drones, acting as a relay between the people of the ground and the fixed infrastructure.
The other challenge concerns current drone regulations that forbid flying drones without a certified pilot, meaning that autonomous flying is prohibited. So, we are waiting for the regulatory aspect to evolve as well. It is analogous to the idea of autonomous cars; they exist, but they are not allowed to be driven in Europe yet.
Another more technical limitation for the moment is the energy autonomy of drones. Currently, they have a flight time of around 30 minutes, which is not sustainable as a solution for a potential product.

Now, regarding the distributed implementation of wireless networks, the main issue is that we need enough consensus among the lead players, in order to change the way we design wireless networks, since it affects fundamentally everything. It is not an add-on that we can propose for a network, it changes radically the way we deploy it. We would need to wait for a political momentum behind this idea, but it is uncertain if it is ever going to happen since there is too much at stake for all the parties already involved. Now things have been changing because of open source, which is acting as a distraction, facilitating small players to enter this market with their own ideas, so maybe things will be changing more rapidly.

Q. And your conclusion remarks at the end of a 5-year ERC project?

DG. First of all, I encourage all my colleagues to try and pursue the ERC funding, which is a great opportunity to do research with very little constraints. For example, what seemed as a small idea in my proposal turned out to be the central point. There is no consortium like in normal EU projects, often lead by industry players, so if you want to change the course of your project you won't need to ask permission. So, the ERC gives you a lot of freedom, but I am sure they already know!
Through this project, I am convinced that the key opportunities in our area, namely communication systems, is the interaction between robotics and communications, an aspect I didn't consider before submitting this ERC project. Drones is an example, but it is not the only one. In the future, it is clear that there will be more robots, used in homes, stores, factories, etc. and we will have teams of robots, needing to work together, communicate and interact. Thus, their communication has to be as efficient as possible, meaning that we need to design proper connection protocols and decision making algorithms at the level of the robot, taking into account the limitations of the communications protocols. For example, if we need robots making decisions based on camera/sensors/gps inputs, like in the case of autonomous cars, it will not be possible to transmit such a large amount of collected data. We would have to be able to identify the essential information and choose which data to share with the network. These challenges, at the interface between the decision making protocols and communication protocols, are going to be very exciting to solve.

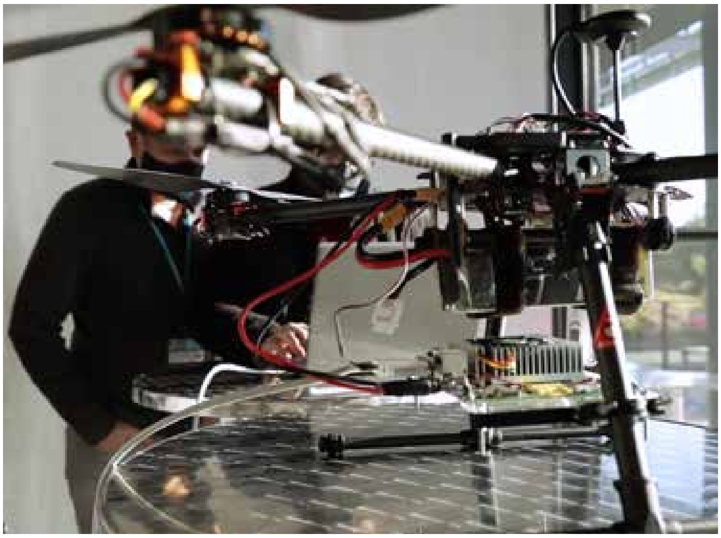
Adlen Ksentini
Professor, Communication Systems Department

How can drones play an essential role for smart cities safety?

As smart cities are emerging, connectivity needs are increasing dramatically in various scenarios of citizens everyday lives.

Unmanned Aerial Vehicles (UAV), largely known as drones, are expected to play a key role in smart cities environment, providing services for traffic management, public safety, situation awareness and connectivity during large crowd events.

EURECOM's professor Adlen Ksentini, expert in mobile and wireless networks, explains the potential role of drones within smart cities. Currently, EURECOM is a key partner of the European project 5G!Drones, demonstrating that 5G is crucial in supporting challenging use cases that could put pressure on network resources, such as low-latency and reliable communication, massive number of connections and high bandwidth requirements, simultaneously.



Q. Could you describe the context of drones use within smart cities?

AK: In the context of smart cities, drones could be used for infrastructure surveillance, public safety and even city ressources optimisation management like waste collection, with the help of IoT devices. For example, monitoring the evolution of a public building construction is essential for public safety. Drones are equipped with 3D cameras, that could create a 3D map of the building and the surrounding area, in order to identify potential issues. These drones are typically connected to a mobile network, operating in areas where WiFi can not cover. The 3D cameras on drones send high data rate stream, which requires a good connectivity provided by 5G. Moreover, the necessary computations for the video reconstruction from data streamed by drones, can be treated locally at the edge of the network of the city, without the need of distant centralised computing solutions.

Another example is to use drones for monitoring the security of infrastructures, by carrying out regular surveillance patrols by drones. This use case is already applied at the Campus SophiaTech in Sophia Antipolis France. All data are sent to the edge of the network, where they are treated with Machine Learning techniques, gaining enormously in latency.

Finally, drones could be placed at the entrance of stadiums, collecting video streaming and sending the collected data to a local server at the edge of the network. This local server would have facial detection algorithms implemented, in order to ensure safety in the sports event. For example, this scenario could address the issue of violence in stadiums by monitoring the crowd flow. Drones could be placed in different positions and can move in order to assist police forces. It is important to note that these data, which are privacy sensitive, would be treated locally at the edge of the network, without the need of central distant servers, addressing privacy protection issues.

Q. What about EURECOM's role and who would be the major stakeholders for drone deployment is smart cities?

AK: EURECOM is contributing a 5G connectivity testing platform, based on OpenAirInterface (OAI), to verify and test that vertical markets are ready to propose their UAV solutions. Our research focuses on multi-access edge computing (MEC) and network virtualisation (NFV), a domain which essentially studies how networks can be self-managed and self-controlled in a virtual environment. We are working on a large range of network use cases, related to orchestration and management of virtual resources using Machine Learning, and Radio Access Networks in terms of virtualisation.

For the smart cities related use cases, we collaborate with the drone operators and owners. They include the video equipment, as well as they the software to make the drone flying. Network operators would provide the connectivity and municipalities would contribute with the infrastructure in terms of hardware and possibly with 5G private networks for specific use cases in smart cities. Finally, vertical markets would use drones for their respective use cases of interest.

Q. What are the challenges for the future regarding drones deployment in actual cities?

AK: It has been 2 years since the beginning of the project 5G!Drones and we are now ready and capable to provide a robust 5G connectivity, to support services related to smart cities safety. We have also already developed the necessary algorithms to be able to host these solutions at the edge of the network. As a next step, drone deployment is currently taking place and the EU Commission is starting to create rules for regulating drone fly, much similar to aerial planification management. Hence, in the next decade drones will be part of our lives. Many verticals are now thinking about using drones as well, e.g. Amazon is considering to transport packages from one place to another, to avoid using trucks. There is also a potential use case to test, concerning people at remote locations needing drug delivery with drones in case of emergency.

Moreover, drones could be essential for mission-critical applications like natural disasters and rescuing people that are found in out-of-reach areas. We are currently working on a project to achieve video stream transmission from drones directly to virtual reality glasses, allowing critical-mission leaders to monitor situation more efficiently.

Drones are still somehow controversial but the social perception is starting to change. Especially after the first lockdown, where drones were used to monitor crowd circulation making sure people respected physical distance measures. Technological prerequisites for UAV use cases in real cities are already achieved. Now, the next step is for society to accept drones and authorities to create a safe regulatory framework for drones to operate.

REFERENCES

http://www.ercperfume.org

PERFUME Publications
http://www.ercperfume.org/pubs/

How can smart territories shape the future of automated driving?

Jérôme Härri
Professor, Communication Systems Department

With the rapid evolution of automotive technology enhanced by advanced AI, having autonomous vehicles actively present in our life is getting closer to reality. However, there are still several issues to be addressed, in order to ensure that AI-enhanced self-driving cars could eventually handle all aspects of driving safely. One of the most complicated scenarios that automated cars have to face is successfully driving through complex traffic situations, such as not-signalised intersections, roundabouts or mixed traffic zones.

In fact, safely crossing these unpredictable traffic situations requires complex decisions for which AI-enhanced automated cars would need to be trained very cautiously. Avoiding these complex traffic situations altogether for AI-enhanced self-driving vehicles would be an inefficient solution, since it would increase pollution and transportation cost, a contradictory concept to the purpose of automated driving.

EURECOM's Professor Jérôme Härri, expert in connected cooperative automated mobility (CCAM), explains how smart territories could play a crucial role in addressing the automated vehicles safe circulation through realistic environments and complex traffic situations such as roundabouts. This work has been in collaboration with his PhD student Duncan Deveaux, who has recently successfully defended his thesis.

Q. What is the state of the art for automated driving cars driving in smart cities?

JH: Automated driving corresponds to autonomous vehicles equipped with technology like AI, sensors, cameras, and exchanging information over advanced communication technologies with other vehicles and road infrastructures. This technology enables vehicles to drive autonomously without or with minor human interventions. Currently, looking at automated driving, we observe that there is a big mismatch between vehicles, that are becoming smarter in autonomous driving and the road infrastructure, which is not designed to support the AI embedded in autonomous vehicles. We can see that by observing that SAE level 5 (full autonomy) vehicles are already capable of self-driving in well-defined and protected zones, whereas SAE level 3 (partial autonomy) are currently expected to drive in more general road infrastructures (for comparison, the Tesla 'autopilot' is considered SAE level 2). Reaching full automation is not going to be achieved by the sole efforts of the automotive industry and smart territories are expected to take a leading role. Road infrastructure should not be limited to providing signalling, but should also be equipped with AI technologies supporting autonomous driving. Considering a roundabout for example, an autonomous vehicle needs to decide on priority by following a yield sign. Its decision is based on its assessment if inbound vehicles are going to exit the roundabout or not, which is a complex task even for a human, and which is specific to each roundabout. It is therefore an opportunity for smart territories, which built these roundabouts and have knowledge of traffic conditions to provide assistance with adequate and precise knowledge to automated vehicle on the best strategy to drive safely through such complex traffic

Q. What are your collaborations and partnerships on this topic?

JH: We have a partnership with TOYOTA InfoTech Labs, USA for the concept of knowledge networking between vehicles and road infrastructures in vehicular networks. Also, this work is part of an EU project IntellIoT, one of which objective is to build a decentralised secured trustworthy IoT framework for developing knowledge networking as a service and potentially paving the way towards a decentralised knowledge marketplace.

There is an entire ecosystem related to both partnerships, spanning from automotive, manufacturing, farming or even health industries. However, we believe that smart territories should also be integrated into this ecosystem. Indeed, our territory-centric approach aims to remove the pressure from the automotive industry to provide autonomous cars that are equipped a priori with full AI knowledge about driving policies, in all potential complex traffic situations. Instead, smart territories' road infrastructures would play a leading role in providing, potentially selling, such knowledge and orchestrating the AI learning process according to the specific contexts and data owned by smart territories. An aspect to keep in mind is that connected automated mobility is not a public service and providing the required knowledge for safely and efficiently driving autonomously may be part of a promising smart territories' led marketplace.

Q. What is the contribution of EURECOM in these projects and what are the challenges you have yet to overcome ?

JH. We strongly believe that the involvement of smart territories infrastructure with guidelines for safe and efficient driving would be crucial for automated vehicles, particularly for unpredictable and episodic scenarios. Here, AI meets smart territories, where cars would receive knowledge created by smart territories providing instructions on how to safely and efficiently drive through complex traffic situations.
In our study, we take the particular case of roundabout representing a complex traffic situation. We define risk reasoning in roundabout in the form of exit probability as well as time to collision, and developed AI models capable of assessing such risk. We then analyse the characteristics of such AI models, to assess if they are specific to each roundabout or if various roundabouts with similar contexts (e.g. number of exits, size, traffic density, number of lanes) could use the same AI models. More important, we propose a semantic, uniquely-defining roundabout risk-reasoning AI models and their applicability context, such that models can be referenced, queried, shared and finally used independently by different actors. We observed that there is a significant need for standardisation in this domain, where the lack of globally agreed ontologies defining what a model is and does under which conditions forbids efficient knowledge networking. Accordingly, EURECOM created a framework to create, name, store and share knowledge between vehicles and smart road infrastructures.

In more technical detail, EURECOM's contributions are summarised as:

1. Defining an AI model capable of quantifying roundabout risks and hazards
2. Clustering roundabouts based on their common features, characterising them in terms of similarities in AI risk assessment
3. Extracting the semantics out of the features similarities in roundabouts, and defining them as parameters uniquely identifying a risk reasoning AI model and its applicability context
4. Implementing a prototype in EURECOM's IoT platform with all the learning, storing and sharing mechanisms, using a vehicular/driving simulators capable of precisely modelling various roundabouts with vehicles equipped with controllable sensors.

Another point is that EURECOM did not only focus on how to identify and disseminate knowledge created by different stakeholders, but also developed an orchestration service framework for decentralised knowledge creation, where each vehicle could collaboratively participate to knowledge creation. Decentralised AI mechanisms such as Federated Machine Learning are promising technologies to benefit for the wisdom of the crowd to train AI models, but their efficiency strongly depends on the dynamic selection of the most appropriate 'crowd'. Benefiting from the defined AI semantics for roundabouts, the proposed framework allows to identify which vehicles could have the most efficient roundabout environment to train a particular AI model.

However, a serious limitation yet to overcome is the system's security. It is important to make sure that smart territories are attack proof, so nobody could potentially inject fake knowledge in the system. So the next step, is to secure knowledge and trace where it came from. To address this issue, we are now starting to use Blockchain-based Distributed Ledger Technologies, which may log every knowledge-related transaction made by any vehicle. However, adapting such technologies to vehicular networks is not straightforward.

Q. What do you think is the future direction of this work ?

JH: A knowledge marketplace. Once again, I think that we should not rely only in automotive industry to train autonomous cars to learn how to drive in all conditions. It is a very important aspect but it should not be the only one available. We need to engage territories to deploy models of risk assessment in their road networks and make such knowledge available to automated driving cars. Large AI stakeholders as well as the automotive industries will probably own knowledge related to autonomous driving. Whoever owns knowledge controls it eventually. We foresee unconstrained access to knowledge in future networks bearing similar societal challenges, as it was the case for information in the pre-Internet era.

In this context, we would like to provide an ecosystem that motivates the territories to take an active role in knowledge creation and sharing for connected automated mobility. After all, they already hold a key position, owning most of the required data and accordingly should not be kept away from this promising future revolution. Benefiting from next generation IoT networks integrating innovations such as private 5G, decentralised AI and distributed ledger technologies, smart territories could create knowledge related to road infrastructure potentially integrating it in a marketplace available to smart vehicles. In turn, the automotive industry could rely on smart territories to provide complementary knowledge supporting more efficient driving in specific contexts. It is a mutual benefit: smart territories could profit from innovative sources of income by providing high quality knowledge to a knowledge marketplace, whereas smart vehicles could have unconstrained access to various sources of applicable knowledge required for autonomous driving. For example, as a first step, we could imagine smart territories using traffic data from their large-scale traffic or security surveillance system, or equipping city vehicles and signalised intersections with AI and private 5G capabilities to generalise knowledge creation as we did on a reduced set of roundabouts.

REFERENCES

https://intelliot.eu/

The start of a new era for EURECOM

Sebastien Wagner,
Engineer Head of Standardization

Today, EURECOM has a full-time expert dedicated to the standardization process for Communication Systems. It is likely that in the future the two other departments, Data Science and Digital Security, will also benefit from this new expertise.

Sebastian Wagner, you have been appointed as the head of Standardization, a job that is actually new at EURECOM. Can you describe what it is about?

I am in charge of the standardization activity at EURECOM, especially for the Communication Systems department. Which means I have three main objectives. First, I have to collaborate with EURECOM researchers so that I can connect their research to current or future standard activities.

Second, I need to contribute to different standards organizations and defend our ideas during the meetings. And my third objective is to generate and manage intellectual property for the department.

Your work experience obviously made you the right person for the job. What is your background?

As a researcher, I've worked on several occasions in standardization processes, especially 3GPP, the 3rd Generation Partnership Project that gathers standards organizations developing protocols and specifications for mobile telecommunications. I've been involved in 4G and 3GPP since the beginning of my PhD in 2008 at EURECOM.

I have developed expertise in wireless communications through innovation and standardization, worked on innovation tools and the OpenAirInterface (OAI) platform.

The OAI platform is EURECOM's flagship in the development of building blocks of mobile communications standards. It must have been the perfect tool to become an expert in the field!

Yes! It was the best possible experience for this job at EURECOM. The objective with OAI is to operate future terminals and infrastructures in open source and have more industrials use the platform.
During my post-doc at EURECOM, I had the opportunity to develop and implement physical layer algorithms for OAI - including synchronization and cell search, random access procedure, advanced uplink and downlink receive algorithms. Then, after 3 years at Intel, I spent the last 4 years at TCL Communication - the Chinese electronics company - as a wireless system expert involved in innovation and standardization. Which means this job at EURECOM is coming after more than 10 years of work in the 3GPP field.

What is your goal and your strategy for developing standardization within the Communications systems department?

It all started a few years ago with Raymond Knopp who started the standardization activity. Today, we want to achieve three goals. First, the 3GPP ecosystem should help us work on problems of immediate practical value and better understand the assumptions of the industry. Then, being in direct interaction with this ecosystem will make it easier for our ideas to be understood and accepted. And of course, it is potentially a major source of funding since researchers can make money from their patents if they protect their ideas.

As for our strategy, it is based on the right balance between long-term and short-term research - in other words fundamental and practical research. For the long-term, the idea is to contribute to core technology that will shape 6G and beyond, which requires a good knowledge of the limitations of the current technology. Whereas short term means practical research within the constraints of the current and upcoming releases of the standard. EURECOM researchers need to contribute to specific areas of 3GPP, like features of the next release, Rel-18 and promote their ideas. They can actually use our OAI platform to validate them. Short term research requires knowing the standard in detail and some agility to follow the 3GPP process.

I guess the idea is to build a team dedicated to the standardization activity and to work more closely with Qualcomm with which EURECOM has a contract?

Indeed. Raymond and I plan to hire a post-doc by the end of 2021. We will also train young PhD students to the reality of industry so their work could be implemented into the standard. Our team will probably get larger in the future since EURECOM's role within 3GPP should be more and more active – especially due to our partnership with Qualcomm that started in 2019. The "Patent Factory Contract" with Qualcomm means we are directly involved in the standardization process. Actually, Qualcomm is the major player in 3GPP. Their expertise greatly helps us promote our ideas within 3GPP – ideas that can turn into standard essential patents. It also helps us develop a sense of where things are going in the future, which is a major advantage!

Apart from 3GPP, we think we could also have an impact on other standards bodies such as ETSI (European Telecommunications Standards Institute) and International Telecommunication Union (ITU-T for 6G) - especially within the Focus Group on Technologies for 2030 - or even the small cell industry.

Isn't it quite uncommon for a research center to be involved in standardization, which is usually taken care of by the industry?

You're right but it already exists! Promoting ideas and being part of standardization processes are part of Fraunhofer's DNA for example, one of Europe's largest application-oriented research institutes. And it's true that EURECOM is far more industry-oriented than the average.

Just look at the number of our partnerships and the OAI platform. To be involved in standardization is what Raymond has in mind for a number of years now. Let's make it clear: Industry always benefits from the results of research institutes. Which means that our work is very significant and that we could play a bigger role in the standardization process along with big companies like Huawei, Ericsson or Nokia. Of course, we'll be a small player compared to them, but we will focus on specific features where we can make an impact. Plus, being more engaged in the process is a good way to enhance intellectual property and create a portfolio at EURECOM.

It seems obvious that IP and standardization activities could also be very effective for the other two departments at EURECOM. What do you think?

It probably could. Researchers in the Data Science and Digital Security departments will certainly start thinking more about standardization when they see some results in the Communications Systems department. However, we will need to have more in-depth knowledge of the standards bodies that are relevant to these topics. In the meantime, I hope our motto "Protect before you publish" will be applied by researchers of these departments.

But our first objective is to build a small team dedicated to standardization in our department. This way, we will improve standardization culture amongst EURECOM researchers and facilitate intellectual property protection in all areas.

> "Monetizing EURECOM intellectual property can provide additional funding for research. This will make the current activity sustainable and profitable for research. In other words, if we protect our ideas that eventually become part of the standard - before publishing articles - then we can monetize them! And this works for both kinds of research: long term and short term. That means less time required for grant application and more time dedicated to research!"

DORA MATZAKOU
Scientific Communication Specialist

EURECOM's investment in a Scientific Communication mission

A scientific communicator has an essential role; making the scientific content and research results accessible to everyone. I really enjoy being the interface between scientist and society and I am glad to have a job that can help researchers communicate their exciting results through dissemination articles, videos and illustrations!

It was certainly not a job description I had come across in job fairs for young PhDs nor did any of my mentors knew that this career path existed after academia. However, as a researcher, communicating science is something you have to do every day and it is certainly a hard task even between fellow experts; especially since it usually entails explaining quite complex concepts. The key to a successful communication is to take a step back and try to go into your audience shoes. Ask yourself what elements they need in order to be able to follow you. I remember several times the disaster in conferences, when people were trying to explain their topic, diving directly into technical details without giving any context.

But what about communicating science to the wider public, like prospective students, future researchers or simply individuals that are passionate about science?

Well, I strongly believe that one of the basic roles of a scientist is to educate society, establishing trust and direct communication with the public. Admittedly, it is quite a laborious mission!

Nonetheless, to illustrate the distance between researchers and society nowadays, lets follow the actual journey of a new scientific achievement to the public, once researchers come up with it. First, scientists have to write a long and complicated journal paper and -if they are lucky- after some review rounds it will be published in a top science journal or magazine. Now, most probably, their article is not accessible to the public for free (well that is a sore issue for another time) and even if it was, it would be very complicated for most readers. In order to gain some visibility to the eyes of the public, their institute might be doing some dissemination activities with articles and videos, diffusing research results in social media and press releases, actions that are essential but unfortunately not yet universal. Most often, the scientific news journey stops here, but yet a third step is necessary to reach the wider public; the attention of general media, unfortunately a quite rare case for research projects.

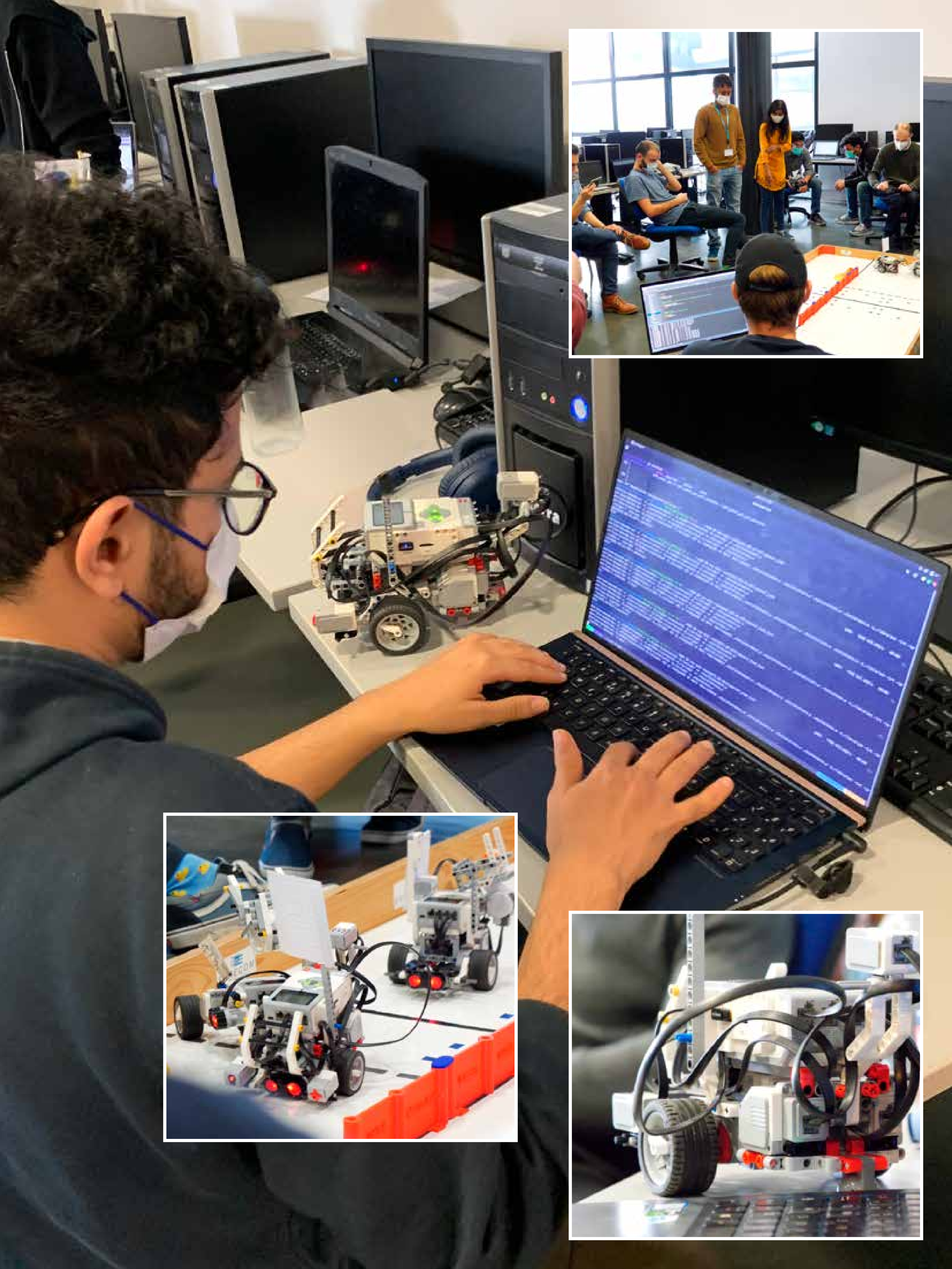
So, why is it important to communicate science after all?

In a nutshell, science communication is an essential part of the research-society ecosystem. Misinformation is one of our largest problems nowadays with deep consequences and there is a huge potential for scientists worldwide to help fight it. Taking the example of the current pandemic lasting for two years now, along with climate change, we understand that it is crucial to communicate scientific truths clearly, in order to address critical situations. We need well-informed leaders and decision makers, in order to shape government policies, ideally based on evidence-based scientific results. In a smaller scale, good scientific communication from research institutes and universities could help educate funding bodies on current research results and future directions, possibly creating a positive loop for funding opportunities.

Also, scientific communication could stimulate collaborations between academic laboratories and industrial partners, as well as attract prospective students and young researchers. Research institutes have the power to establish themselves as the official channels of scientific information flow, that the public can trust and use. A better informed public is a public with critical thinking, participating constructively in community conversations. And who knows, a passionate researcher explaining his most exciting findings, could be the role model that children need to be inspired and become the next generation's scientists!

REFERENCES

Learn about the exciting research at EURECOM in our research blog at https://eurecom-blog.medium.com/



EURECOM GIE MEMBERS

Politecnico di Torino

Aalto University

Technische Universität München

NTNU Norwegian University of Science and Technology

CHALMERS UNIVERSITY OF TECHNOLOGY

CZECH TECHNICAL UNIVERSITY IN PRAGUE

TU WIEN TECHNISCHE UNIVERSITÄT WIEN Vienna | Austria

ITMO UNIVERSITY

LIÈGE université

EDHEC BUSINESS SCHOOL

orange™

NortonLifeLock

SAP®

BMW GROUP

Gouvernement Princier PRINCIPAUTÉ DE MONACO

Institut Mines-Télécom

WANT TO JOIN US?
communication@eurecom.fr

EURECOM GRADUATE SCHOOL AND RESEARCH CENTER IN DIGITAL SCIENCE

WWW.EURECOM.FR

INSTITUT CARNOT
Télécom & Société numérique

EURECOM - CS 50193 - 06904 Sophia Antipolis cedex