TELECOM
Paris

IP PARIS

2021
**Review**
Rétrospective

A year of
research
at Télécom
Paris

Un an de recherche
à Télécom Paris

# 2021

**Review**
Rétrospective

# Table of contents
# Sommaire

# Editorial
## Édito



**Nicolas Glady,** Dean and President

**EN** I am delighted to present A Year of Research at Télécom Paris for the second time, it will enable you to find out about our research activity or explore it further. Télécom Paris, this leading and more than century-old engineering school is based on a tradition of scientific excellence, addressing the major challenges of our time and even pre-empting them. They include embedded systems and the digital economy, and also networks, 5G, the cloud, cyberspace, algorithmics, big data and quantum computing.

A higher education institution under the trusteeship of the French Ministry of the Economy, Finance and Recovery, our research themes are now recognized as topics of digital sovereignty and strategic autonomy. This historical trend is now even more marked, in the wake of recent crises. The Covid pandemic comes to mind, as does geopolitical turmoil, making us rethink our practices and adapt better to the needs of a world where contingencies are becoming the norm.

Our contribution to the study and achievement of breakthroughs in critical areas such as very large digital infrastructures (especially the cloud), data science taken in its broadest meaning, artificial intelligence, the ethics of algorithms, digital trust and cybersecurity, shapes our Raison d'être*. Mindful of humans and their environment, we enshrine these principles in our research activity by developing responsible digital solutions and being aware of environmental issues. You will see that for yourself, as you browse through this review of our scientific output. Just like organizations with a mission, this Raison d'être is a powerful lever for transformation and guides our work for a positive social impact.

Our school needs to mirror changes in the economy and in society, to meet the need for innovation and support for entrepreneurship. Concerning these two aspects, Télécom Paris has reinforced its teams to provide the training and support required in the light of these changes.

Télécom Paris remains an open sky lab where academic excellence combined with a culture of innovation and entrepreneurship enables us to affirm our expertise in the deep tech required for technological and economic developments, now and in the future.

As you read this document prepared for you, you will learn about the many aspects of our research activity which feeds into our training programs, from the engineering cycle to doctorates, including post master's degrees and our executive education offering. I hope you enjoy reading this review!

---

**\* Training, imagination and entrepreneurship to design models, technologies and digital solutions, serving a society and an economy mindful of humans and their environment.**
Former, imaginer et entreprendre pour concevoir des modèles, des technologies et dés solutions numériques, au service d'une société et d'une économie respectueuses de l'humain et de son environnement.

**FR** Je suis très heureux de vous présenter la seconde édition d'Une Année de Recherche à Télécom Paris qui vous permettra de découvrir ou d'approfondir votre connaissance de notre activité de recherche. Grande école d'ingénieurs plus que centenaire, Télécom Paris s'appuie sur sa tradition d'excellence scientifique pour répondre aux grands enjeux de notre temps, et même les anticiper : des systèmes embarqués à l'économie du numérique, en passant par les réseaux, la 5G, le cloud, la cyber, l'algorithmique, le big data ou l'informatique quantique !

École du ministère de l'Économie, des Finances et de la Relance, nos thèmes de recherche sont dorénavant reconnus comme des sujets de souveraineté numérique et d'autonomie stratégique. Cette orientation historique se trouve aujourd'hui renforcée par la succession de crises que nous venons de traverser. Je pense ici particulièrement à l'épidémie de Covid-19 mais aussi aux troubles géopolitiques qui nous amènent à repenser notre pratique pour être plus adaptés aux exigences d'un monde où l'aléa devient la norme.

Contribuer à la réflexion et aux avancées sur des sujets cruciaux comme les très grands équipements numériques (Cloud notamment), la science des données dans son acception la plus générale, mais aussi l'intelligence artificielle, l'éthique des algorithmes, la confiance numérique et la cybersécurité fait donc partie de notre Raison d'être*. Respectueux de l'humain et de son environnement, nous incarnons ces principes dans notre activité de recherche par le développement d'un numérique responsable et conscient des enjeux environnementaux. Vous pourrez le constater par vous-même en parcourant ce panorama de notre production scientifique. Cette Raison d'être qui, à l'instar des sociétés à mission, est un levier puissant de transformation, oriente notre action pour un impact social positif.

Notre école doit accompagner les évolutions de l'économie et de la société pour répondre aussi aux besoins d'innovation et de soutien à l'entrepreneuriat. Sur ces deux dernières dimensions, Télécom Paris a renforcé ses équipes pour offrir la formation et le soutien nécessaire à ces évolutions.

Télécom Paris demeure un laboratoire à ciel ouvert où l'excellence académique, alliée à une culture de l'innovation et de l'entrepreneuriat, nous permet d'affirmer notre expertise dans les deep tech nécessaire aux développements technologiques et économiques présents et à venir.

En parcourant ce document élaboré pour vous, vous pourrez appréhender les multiples aspects de notre activité de recherche qui nourrit nos enseignements, du cycle ingénieur au doctorat en passant par les Mastères Spécialisés et notre offre d'executive education. Bonne lecture !

# Key figures

## Télécom Paris en chiffres

**Disclaimer**

EN This document concerns the research activity of our teams in 2021. In order to obtain the most representative figures possible, unless otherwise stated, the statistics shown on these pages and on the Team pages are taken from 2021.

**Avertissement**

FR Ce document concerne l'activité de recherche de nos équipes en 2021. Afin d'obtenir les chiffres les plus représentatifs possibles, les indicateurs sur cette page et sur les pages Équipes sont pris sur l'année civile 2021, sauf indication contraire.
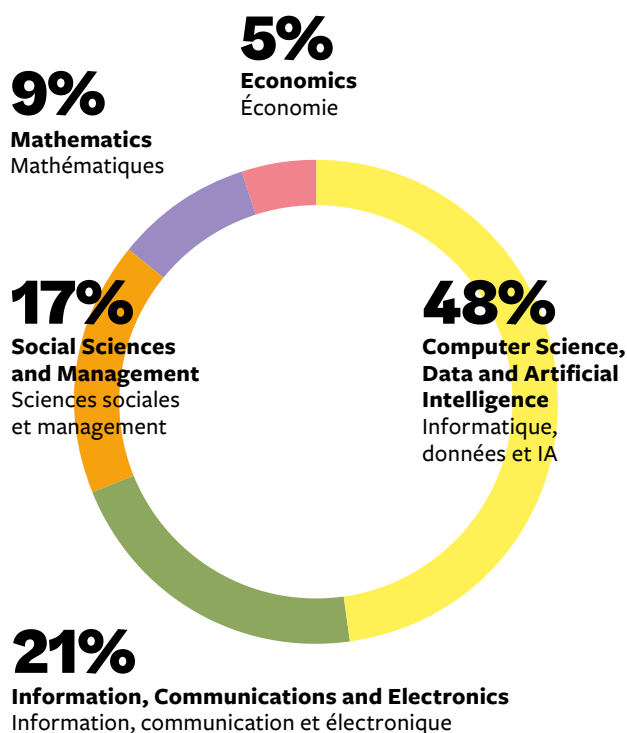
## Staff
## Personnel

**154**
**Permanent Researchers and Faculty Members**
Chercheurs et enseignants-chercheurs permanents

**242**
**PhD Students and Postdocs**
Doctorants et post-doctorants

**83**
**R&D Engineers**
Ingénieurs recherche et développement

**26**
**Administrative Staff**
Personnel administratif

**9%**
**Mathematics**
Mathématiques

**5%**
**Economics**
Économie

**17%**
**Social Sciences and Management**
Sciences sociales et management

**48%**
**Computer Science, Data and Artificial Intelligence**
Informatique, données et IA

**21%**
**Information, Communications and Electronics**
Information, communication et électronique

Breakdown by Institut Polytechnique de Paris Research and Education Departments / Répartition par Départements d'enseignement et recherche de l'Institut Polytechnique de Paris

**505**
Total Research Staff
Effectif total de la recherche

**18**
Research Teams
Équipes de recherche

# Scientific results
## Résultats scientifiques

**178**
Journal Articles
Articles dans
des journaux*

**231**
Conference Papers
Articles dans
des conférences
internationales*

**32**
Book Chapters and
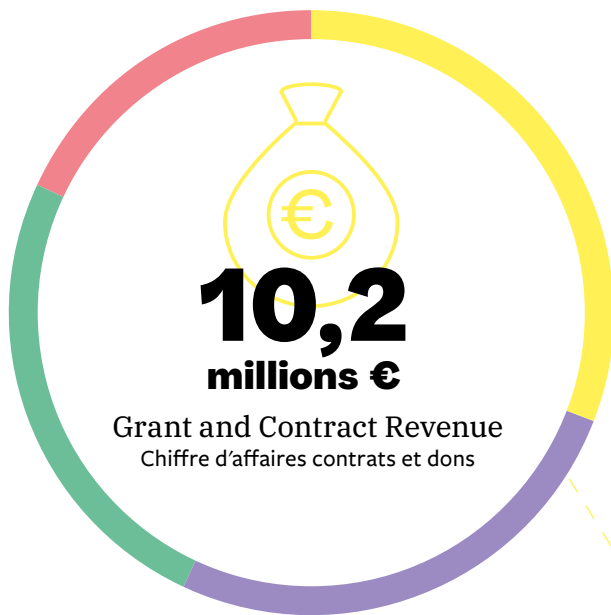Proceedings
Livres et chapitres
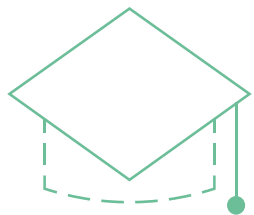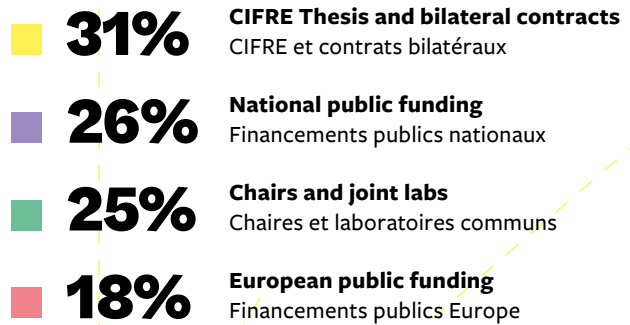de livres*

**49**
PhD Defenses
Thèses soutenues

**64**
Other publications
Autres publications*

*The figures have been calculated with aggregated data for the 2020-2021 period

**10,2**
**millions €**
Grant and Contract Revenue
Chiffre d'affaires contrats et dons

# Research Funding
## Financement de la recherche

**31%** CIFRE Thesis and bilateral contracts
CIFRE et contrats bilatéraux

**26%** National public funding
Financements publics nationaux

**25%** Chairs and joint labs
Chaires et laboratoires communs

**18%** European public funding
Financements publics Europe

# ERC Grants
## Bourses ERC

▸ telecom-paris.fr/erc

Since the beginning of the ERC Program /
Depuis la création du programme ERC

**Mansoor Yousefi**
ERC Starting Grant 2018

**Michèle Wigger**
ERC Starting Grant 2016

**Yanlei Diao**
ERC Consolidator Grant 2016

**Alexandre Gramfort**
ERC Starting Grant 2015

# Innovation and transfer
## Innovation et transfert

5-year consolidated figures / Chiffres consolidés sur 5 ans

Spinoff creation
Spinoff créées
**2**

Maturing projects
Projets en maturation
**24**

Pending patents
Brevets déposés
**79**

Registered software
Logiciels déposés
**14**

Registered technical know-how
Savoir-faire protégé
**1**

# Highlights

## —— Temps forts

### 15 September 2020

#### Hi! PARIS launch

HEC Paris and the Institut Polytechnique de Paris (IP Paris) launched Hi! PARIS, a new interdisciplinary research and teaching center dedicated to data science and human-centered artificial intelligence. It draws on the 300 researchers and infrastructure of IP Paris and HEC Paris to address the main challenges related to technological transformation and its impact on businesses and society.

Hi! PARIS brings together world-class researchers, professors and PhD students to create Europe's first interdisciplinary and interinstitutional center, combining education, research and innovation, and dedicated to pushing back the boundaries of AI and data science. The Center, which aims to become a world leader in the field within five years, is being supported by global companies such as Kering, Rexel, TotalEnergies, L'Oréal and Capgemini.

▸ See also IP Paris page 90.

### 19 November 2020

#### IP Paris Electrical Engineering Artificial Intelligence Day

During this one-day event, the current research in artificial intelligence and electrical engineering has been reviewed through plenary talks, given by leading scientists, and illustrated with examples taken from well-known laboratories as well as from the work of IP Paris teams.

The seminar concentrated on the interplay between AI and the electrical engineering disciplines (AI and information theory, AI and communications and AI in electronics and optics), and showed how this interdependency has the potential for decisive innovations.

Introduced by Bruno Thedrez, Head of the Information, Communications and Electronics (ICE) Department, Institut Polytechnique de Paris, the talks addressed questions such as implementation, speed, energy, and examined the foundations of AI as well as its theoretical developments in the Electrical Engineering disciplinary fields.

### 24 November 2020

#### Grand Prix IMT-Académie des Sciences

Gaël Richard, Professor, Head of the Image, Data, Signal department at Télécom Paris and Executive Director of Hi! PARIS, was awarded the 2020 Grand Prix IMT-Académie des Sciences.

His research work, anchored in the field of digital transformation sciences and technologies, focuses on the analysis, transformation and recognition of sound signals, and brings into play advanced principles of statistical signal processing, optimization and machine learning.

This annual prize, established by IMT in partnership with Fondation Mines-Télécom, rewards a scientist for his contribution to a scientific breakthrough in the field of Digital transformation in industry, Engineering for the energy and the environment or Materials and manufacturing. This body of work is recognized as serving progress and sustainable economy for the industry or the private sector.

### 15 December 2020

#### Inauguration of the NoRDF Research Chair

Inauguration of the "NoRDF: Modeling and Extracting Complex Information from Natural Language Text" chair with EDF, Schlumberger, Converteo and Groupe BPCE as well as the French Defense Innovation Agency (AID). The NoRDF research project, led by Fabian Suchanek and co-led by Chloé Clavel stems from the French National Program for AI and aims at modeling and extracting complex information from natural language texts. Its goal is to enrich the knowledge bases used by artificial intelligence systems with events, causes, conditions, priors, histories, negations, and beliefs. In particular, this project will study the expression of feelings.

▸ See also DIG team, page 50.

# January 2021

## Quantum communication infrastructure

Orange and its partners in the Paris Region Quantum Communication Infrastructure project announced the deployment of the largest quantum communication network in France. This is the longest fiber ever deployed in France for a quantum communication network, with a length of 43 kilometers, linking Châtillon (Hauts de Seine), the Plateau de Saclay (Essonne) and, later, Paris. The network security is based on a quantum key distribution system. The LTCI is one of the three academic laboratories involved in the project.

## 21 January 2021

### New prize for Antonio Casilli's book

The book by Antonio Casilli (SID) *En attendant les robots, enquête sur le travail du clic* (Waiting for Robots: An inquiry into click work, Seuil) has been awarded the Prix de la Fondation Colbert. This prize aims to distinguish economic, social and cultural projects that are both part of Colbert's legacy and the start of a new way of innovating in France. This is the third prize for this book.

## 08 February 2021

### Operational AI Ethics

"Operational AI Ethics" is a program conducted by an interdisciplinary team grouping six academic fields (applied math, statistics, computer science, economics, law and sociology) around the ethical issues raised by artificial intelligence.
A newsletter is launched to bring together the latest articles on the initiative's blog written by researchers from Télécom Paris, their publications in ethical AI, news on the topics covered by the research program, and upcoming events on the subject.

## 23 April 2021

### A World First: privacy in free-space optical communication

In a new publication in Nature Communications, a group of international researchers, including Olivier Spitz and Frédéric Grillot of Télécom Paris, have tackled the question of privacy in free-space optical communication with the help of two uni-directionally coupled quantum cascade lasers operating in the chaotic regime. A proof-of-concept of a private communication link operating free-space in the mid-infrared window was achieved, a world first, opening promising strategies for physical encryption and decryption of a digital message.
In November 2021, Olivier Spitz was one of the three winners of the DGA (French Directorate General of Armaments) 2021 thesis prize. In December, Frédéric Grillot and Olivier Spitz's work was selected by the Optical Society of America as one of the main topics of the year 2021.

▸ **O. Spitz**, A. Herdt, J. Wu, G. Maisons, M. Carras, C.-W. Wong, *et al.*, "Private communication with quantum cascade laser photonic chaos," *Nature Communications*, vol. 12, no. 1, p. 3327, Jun. 2021.

### Télécom Paris' expertise in 5G recognized

The expertise of Télécom Paris professors in the 5G domain is recognized, from expert committees to the general media.
Laura Draetta, a sociologist of the environment and sustainable development, and a specialist in technological controversies related to digital innovations, contributed to the work of the group of experts set up by the Agence Nationale de SEcurité Sanitaire (ANSES) to assess the risks associated with the deployment of 5G.

Joe Wiart (head of the Modeling, Characterization and Control of Electromagnetic Wave Exposure [C2M] chair) has spoken regularly in the media about his work on health risks of electromagnetic waves.

Alain Sibille, Marceau Coupechoux, Gérard Memmi, Maya Bacache and Marc Bourreau, among others, have also been invited in the media as experts in 5G and next generation networks, to discuss issues related to technological sovereignty, security and privacy, environment and energy consumption, uses and economic value.
▸ telecom-paris.fr/news/focus/actus-5g
▸ chairec2m.wp.imt.fr

# Highlights

## 21 May 2021

### FARO and CERES, CIEDS IP Paris projects

Two projects, FARO and CERES, have been accepted at IP Paris as part of CIEDS (*Interdisciplinary Centre for Defence and Security*).

The FARO project (Algorithmic Foundations of Robot Swarms) aims to define new algorithms (model and data driven environments) and new validation methods to ensure the proper functioning of drones in swarms, on land, in the air, at sea, conducting joint operations.
The CERES project (Cybersecurity Evaluation in Realistic EnvironmentS) aims to develop over the long term a methodological framework and the associated tools for the study of complex cybersecurity systems with different security features (specification of security properties, deployment of mechanisms, verification of mechanisms efficiency).

## 30 June 2021

### Annual research day

This annual day is an opportunity for teams to exchange views, through keynotes and demonstrations. Among the topics discussed this year, the concept of age-of-information aware caching updating should be mentioned (Philippe Ciblat, ComNum), as well as talks on the economic analysis of markets for technology and for inventions (Laurie Ciaramella, ECOGE), and works on the estimation and optimization of the power consumption of an analog-to-feature converter (Olivier Fercoq, S²A).
In addition to specific research works, both the Connected Cars and Cyber Security (C3S) research chair and the Intellectual Property and Technology Transfer support team were given a slot to present their activities.
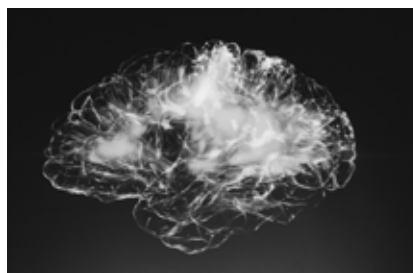


## 15 June 2021

### Launch of TALia, a joint laboratory on natural language processing

Télécom Paris and Onepoint have joined forces to create TALia, a joint research laboratory dedicated to natural language processing with one ambition: to humanize relations with machines.
TALia combines the business and scientific expertise of onepoint employees with the research capabilities of Télécom Paris on a subject of shared interest. It will conduct work on the interaction between machines and humans with a focus on creating solutions capable of understanding natural language integrated in applications found in our daily lives: speech synthesis, spell checker, word prediction on smartphones, and the analysis of opinions and emotions. The laboratory has programmed its work over a five-year cycle and aims to support fout to five theses on this topic.

## 13 July 2021

### National Order of the Legion of Honour



Isabelle Bloch, Full professor at Télécom Paris (expert in mathematical modeling of spatial relations and spatial reasoning, with image processing as a main field of application), promoted to the rank of Knight of the National Order of the Legion of Honour (*Chevalier de la Légion d'honneur*).

## Progress in university rankings

In the 2021 edition of the QS World University Rankings by Subject, Télécom Paris has risen from rank 101-150 to 64th in the world (4th in France and 19th in Europe) in the Computer Science category. Télécom Paris has also progressed in Mathematics (from rank 401-450 to rank 301-350) and Engineering Technology (from 270 to 218) categories. On a national basis, Télécom Paris was ranked 1st Engineering school in digital, computer science and mathematics, and 5th among the "Grandes écoles d'excellence".
These rankings reflect the school's excellence, confirming our position as France's leading engineering school for digital technology. They also reflect the validity of the Télécom Paris model, which trains engineers and PhDs in the theoretical, technological and sociological aspects of digital technology using a unique trans-disciplinary approach.
As a proof of the excellent results of its engineering schools, and for its first entry in the QS rankings, IP Paris has been ranked 49th in QS Global World Rankings 2022, and 15th for the specific Graduate Employability Ranking.

## 25 October 2021

### DSP Workshop in Optical Fiber Communication

Télécom Paris and Aston University organized a one-day online Workshop on Digital Signal Processing in Optical Fiber Communication on October 25, 2021. The workshop brought together researchers from the fields of optical fiber transmission, digital signal processing (DSP) and neural networks, to discuss advances in digital communication over optical fiber. Distinguished speakers from academia and industry provided an overview of topics of interest, and identified important open problems. The workshop was funded by two MSCA ITN-EID European projects REAL-NET and FONTE, where Télécom Paris is a partner.
The workshop received substantial interest from the community, with 140 registrations and a peak-time attendance of 105. Participants attended from the European countries (Denmark, France, Germany, Italy, etc.), Brazil, Canada, China, India, Pakistan, Tunisia, Kazakhstan, UK and the United States.

# 17 November 2021

## A new research team in cybersecurity and cryptography

C², for Cybersecurity and Cryptography, is a new research team of the LTCI. Thanks to the merger of former CCN and MC² teams, and the recruitment of two experts in cryptography, C² strengthens the expertise in the domain of digital infrastructures security. Its research activities span the following domains, including their interactions: cybersecurity, network security, cryptography and privacy, discrete mathematics and fundamental research.

# 18 November 2021

## Data protection impact assessment: the case of connected cars

The 22nd edition of the Values and Policies of Personal Information (VP-IP) chair was co-animated with the Connected Cars and Cyber Security (C3S) chair, in order to present a joint research report. A panel brought together a lawyer at the French Data Protection Authority (CNIL), an expert in cyberphysical and IoT systems, the data protection officer at Renault Group, and a policy and regulatory affairs manager at Nokia.

The report, written by Claire Levallois-Barth and Jonathan Keller in French, addresses the interaction between the personal data protection impact assessment resulting from a legal obligation imposed by the GDPR and the operational needs of the automotive sector. Indeed, data flows in the automotive sector have become more complex, as the connectivity of the vehicle transforms it into an informational platform raising new issues, especially regarding the identification of interested parties and the regulation of personal data sharing.

# 24 November 2021

## SEIDO Lab

By renewing the SEIDO Lab agreement, Télécom Paris and EDF, with its EDF Lab Paris-Saclay research center, are achieving a new level of collaboration in the field of digital technology applied to the energy sector. Cybersecurity and digital trust will be at the heart of a joint research program which will also address IoT and cryptography issues around the Blockchain.



# 25 November 2021



## Our young talents awarded

56 donors took part in the sixth Télécom Paris and Fondation Mines-Télécom gala dinner: this was an opportunity to contribute to the financing of four research projects linked to the digital transformation.

From left to right:

▸ **Stéphane Lathuilière** (Designing an AI that adapts its knowledge dynamically),

▸ **Laurie Ciaramella** (The financing of young innovative companies),

▸ **Jan Gugenheimer** (Ubiquitous mixed reality: building mixed reality technology to fit into the fabric of our daily lives),

▸ **Kyriaki Niotaki** (Power autonomy for the Internet of Things applications).

# Industrial Chairs

Télécom Paris runs an active training and research program in partnership with companies to proactively support and enhance innovation with the help of Fondation Mines-Télécom. Within its Chairs, a multidisciplinary research approach contributes to guide companies, citizens and public authorities. Here is a selection of notable activities in 2021.



## Modeling, Characterization and Control of Exposures

The **C2M** Chair contributes to the characterization, analysis and modeling of exposure to electromagnetic waves induced by telecommunications systems and networks. It brings together Télécom Paris and IMT Atlantique researchers on five research areas: Numerical and experimental dosimetry; Statistical modeling of exposure; Stochastic dosimetry; Influence of network technologies and architectures on exposure, standardization of exposure assessment methods; Parameters governing the social perception of risk. The Chair is led by Joe Wiart, Professor at Télécom Paris and Chairman of the European Committee for Electrotechnical Standardization (CENELEC) TC106x, in charge of standards relating to human exposure.

### 30 December 2021

Work carried out as part of the Chair on the assessment of the risk of wireless phones on brain tumors in young people has contributed to case-control studies in Australia, Austria, Canada, France, Germany, Greece, India, Israel, Italy, Japan, Korea, the Netherlands, New Zealand and Spain.
G. Castaño-Vinyals et al. "Wireless Phone Use in Childhood and Adolescence and Neuroepithelial Brain Tumours: Results from the International MOBI-Kids Study." *Environment international*, vol. 160 107069.

▸ chairec2m.wp.imt.fr

## Connected Cars and Cyber Security

The **C3S** Chair project addresses issues surrounding connected and autonomous vehicles that crystallize some of the most advanced and complex technical challenges of the digital transformation. Its objective is to develop international-level training and research programs, in close cooperation with its partners. In addition, it focuses on new cybersecurity issues linked to the emergence of new mobility patterns, in five major themes: Cybersecurity, risk analysis and dependability; Protection of data and data flow in real time, cryptography and agility; Identity management, authentication; Resilience by design; Personal data protection and connected vehicles. On the latter, the Chair works closely with the Values and Policies of Personal Information (VP-IP)Chair. Guillaume Duc, Professor and specialist in embedded systems security and Rida Khatoun, Professor and expert in network security and attack detection, lead the Chair. The industrials partners of the Chair are Nokia, Renault, Thales, Valeo and Wavestone.

### 18 November 2021

"Data Protection Impact Assessment: The Case of Connected Cars", co-organized with the **VP-IP** Chair.

▸ www.telecom-paris.fr/c3s
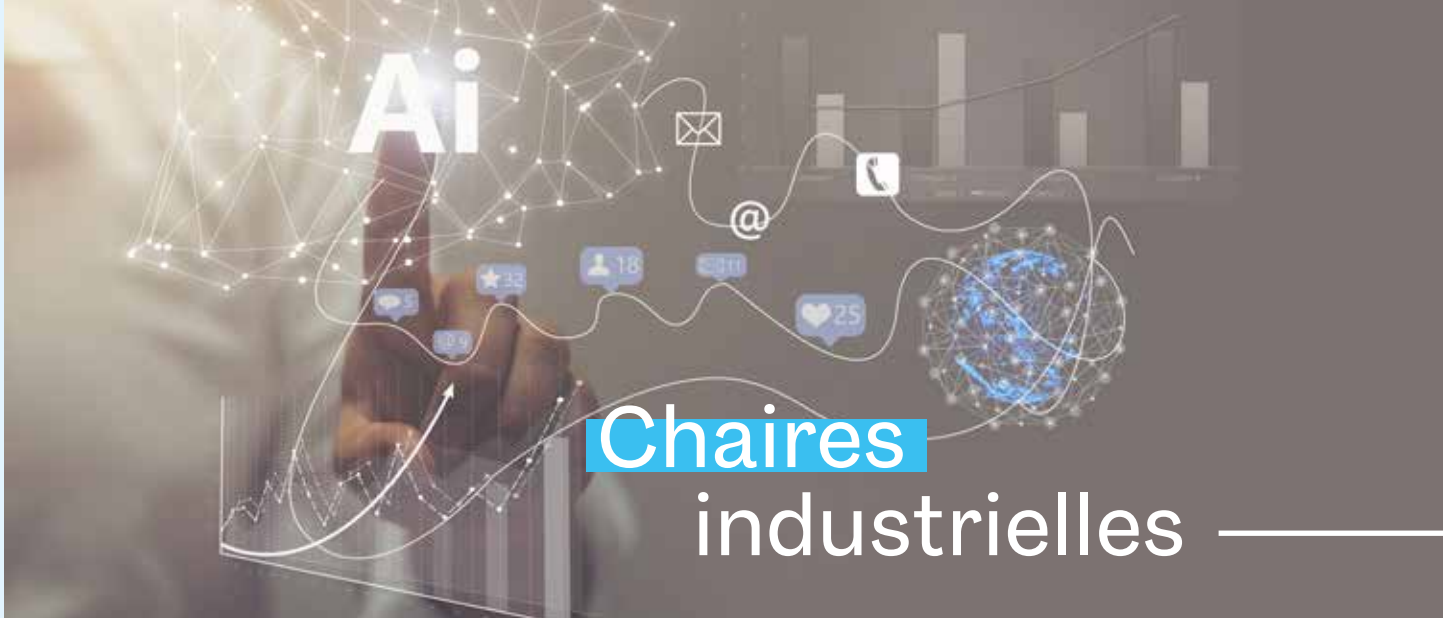
# Chaires industrielles

## Data Science & Artificial Intelligence for Digitalized Industry & Services

The **DSAIDIS** Chair research and training program aims to pool and enrich industrial and academic experiences, to bring out new innovations and to enable sustainable funding of research activities in artificial intelligence and data science on methodological topics that are crucial for applications. It also aims to prepare possible bilateral or multilateral collaborations, and to propose training courses that meet the socio-economic challenges of the field. Led by Florence d'Alché-Buc, Professor in the fields of Computer Science and Applied Mathematics, its research work explores four themes: Building predictive analytics on time series and data streams; Exploiting large scale, heterogeneous, partially labeled data; Machine learning for trusted and robust decision; Learning through interactions with the environment. Five of its research projects were presented at the last NeurIPS conference (2021), resulting in 16 publications in the last three years. The Chair's industry partners are Airbus Defence & Space, Engie, Idemia, Safran and Valeo.

### **17** November 2021

Internal workshop with the industrial partners on "Reinforcement learning".

▸ www.telecom-paris.fr/dsaidis

## Explainable Artificial Intelligence for Anti-Money Laundering

The **XAI-4-AML** research Chair addresses artificial intelligence explainability in the context of anti-money laundering and countering the financing of terrorism. Its objective is to develop an optimal framework for the deployment of artificial intelligence in the fight against money laundering and terrorist financing. David Bounie, Professor of Economics, Winston Maxwell, Professor of Law and Technology and Stephan Clémençon, Professor of Applied Mathematics and Statistical Learning, lead the Chair. Its work contributes to the activities of the Télécom Paris "Operational AI Ethics" program. The partners are PwC, ACPR Banque de France, Dataiku and the French National Research Agency (ANR).

### **19** June 2021

Astrid Bertrand, Winston Maxwell and Xavier Vamparys, researchers at XAI-4-AML, explained in an article published in "Le Monde" that artificial intelligence can make anti-money laundering systems, which currently cost more than they earn in Europe, more effective. *Financement du terrorisme: « Mettre l'intelligence artificielle et le partage de données au centre du dispositif de lutte »* (Financing Terrorism: Placing Artificial Intelligence and Data Sharing at the Heart of the Anti-Laundering Process).

▸ xai4aml.org

## Modeling and Extracting Complex Information from Natural Language Text

The scientific project of the **NoRDF** Chair aims to model and extract complex information from natural language text. It is led by Fabian Suchanek, Professor at Télécom Paris, specialist in knowledge bases and ontologies, and co-creator of the famous YAGO knowledge base. The Chair is co-led on the sentiment analysis component by Chloé Clavel, Professor at Télécom Paris and specialist in natural language processing. The partners are EDF, Schlumberger, Converteo and Groupe BPCE as well as the French National Research Agency and the French Defense Innovation Agency.

### **15** December 2020

Inauguration of the Chair: "Modeling and Extracting Complex Information from Natural Language Text".

▸ nordf.telecom-paris.fr

### **25** April 2021

First workshop with the partners on "Overview of the NoRDF Project; Vagueness, Non-Named entities, and Knowledge Representation; Neuro-Symbolic AI for Reasoning; Text Generation and Storytelling".

# International
## Cooperation

The academic vision of a truly worldwide community makes for extraordinary interactions, fueling the creation and dissemination of knowledge. As a key player in the design and promotion of communication and information technologies, Télécom Paris naturally subscribes to this vision. Since the early eighties and the creation of the first research center at Télécom Paris, its researchers have established relations with universities and research centers all over the world. Cooperation takes many forms, including research partnerships, research visits and sabbatical exchanges, joint workshops and seminars, and also joint PhD supervision that can lead to a double degree PhD diploma. International cooperation also gives rise to numerous publications in major scientific journals and conferences and other joint results. We highlight on this page some of our international cooperation projects.

▸ [1] P. Okelmann, **L. Linguaglossa**, F. Geyer, P. Emmerich, and G. Carle, "Adaptive Batching for Fast Packet Processing in Software Routers using Machine Learning," in *2021 IEEE 7th International Conference on Network Softwarization (NetSoft)*, 2021, pp. 206–210.

Source code and scripts: github.com/pogobanane/vpp-testing

▸ [2] **T. Zhang**, L. Linguaglossa, P. Giaccone, **L. Iannone**, and **J. Roberts**, "Performance Benchmarking of State-of-the-Art Software Switches for NFV." 2020.

[2] has been carried out at LINCS and benefited from the support of NewNet@Paris, Cisco's Chair "Networks for the Future" at Télécom Paris.

▸ [3] C. Shelbourne, **L. Linguaglossa**, **T. Zhang**, and A. Lipani, "Inference of Virtual Network Functions' State via Analysis of the CPU Behavior," presented at the *International Teletraffic Congress*, Avignon, France, Aug. 2021.

▸ [4] **T. Zhang**, **H. Qiu, L. Linguaglossa**, W. Cerroni, and P. Giaccone, "NFV Platforms: Taxonomy, Design Choices and Future Challenges," *IEEE Transactions on Network and Service Management*, vol. 18, no. 1, pp. 30–48, 2021.

### Monitoring and AI-based performance optimization for high-speed networks

This strand of research is conducted in collaboration with several European universities (UCL, TUM, University of Würzburg), and with Nokia.

**RMS** research team pursues its exploration of machine learning (ML) and artificial intelligence (AI) techniques for driving networks without human intervention, in this series of papers where platforms and devices are analyzed ([2,4]), and ML techniques are identified and used ([1,3,4]).

The current trend of "network softwarization", fueled by the growing popularity of network function virtualization (NFV), can significantly reduce the maintenance cost of network services, as the life-cycle of software evolves much faster than that of hardware-based solutions. By replacing the purpose-built, expensive, proprietary network equipment with software network functions consolidated on commodity hardware, NFV envisions a shift towards a more agile and open service provisioning paradigm.

In article [1], batching techniques have been successfully implemented to reduce the HW/SW performance gap (for a methodology to compare fairly and comprehensively the performance of software switches, see also article [2]). Researchers propose, evaluate and discuss an algorithm to dynamically allocate batch sizes depending on the traffic condition instead of the classical busy polling approaches. With the help of a large dataset collected over hours of experiments with a real packet processing engine, ML techniques are first developed to find the optimal batching size for different load scenarios. The training models are then deployed within a software router, and the impact of the approach in terms of saved clock cycles is assessed.

Article [3] develops a novel methodology, based on a "CPU fingerprint", to infer the behavior of high-level virtual network function (VNF) applications, such as throughput or packet loss, using CPU data instead of network measurements. In addition to reducing the number of invasive network measurements, another use-case for our methodology is the resource optimization for high-speed VNFs of unknown requirements. For instance, it is possible to leverage the great amount of available data with lightweight machine learning algorithms that could be deployed within the datapath to quickly react to changes in the application state such as packet loss or misbehavior.

Article [4] is a study of existing NFV platform implementations. Researchers give an architectural overview of the components of NFV platforms, they present their taxonomy on existing platforms based on a typical VNF life cycle, propose a collection of critical design choices and survey the solutions adopted by different platforms. Among other issues and challenges, they envision future directions with AI for NFV platform design. Indeed, researchers believe AI will become commonplace in the NFV domain, it is still in its infancy in the telecommunication industry and several challenges yet need to be overcome to thoroughly unleash the potentials of AI in NFV.

## Inclusive, innovative and reflective societies

The **DIALLS project** — "Dialogue and Argumentation for cultural Literacy Learning in Schools" (dialls2020.eu) — was funded between May 2018 and June 2021 by the European Union H2020 research and innovation funding, under the "Inclusive, Innovative and Reflective Societies" theme. The aims of the project were to engage school children in Europe (and beyond) in discussions together, organized in a way that favored awareness of their diverse cultural identities and also the co-construction of the meaning of key European values, such as tolerance, inclusion and empathy. Coordinated by the University of Cambridge (UK), the project consortium comprised universities across nine EU countries, including the UMR CNRS i3 at Télécom Paris (action led by Michael Baker, DR CNRS, and Françoise Détienne, DR CNRS), and Israel (Hebrew University of Jerusalem). Students' group activities, both in the classroom and across classrooms/countries using online discussion tools, were structured around the study of "wordless texts"—picture books containing a narrative sequence of images, involving ethical issues relating to European values. The project led the CNRS-Télécom Paris team to employ two full-time post-doc researchers (Dr. L. Bietti and Dr. G. Pallarès), who had two main responsibilities: (1) coordinating the design of a specialized Internet platform that enabled shared discussion around annotation of still and video images; (2) developing a method for analyzing the online discussions, in terms of levels of interpretation and negotiation of meanings for ethical concepts.

Three chapters written by CNRS-Télécom Paris authors are included in the following book:

▸ F. Maine and M. Vrikki (Eds.) (2021). *Dialogue for Intercultural Understanding: Placing Cultural Literacy at the Heart of Learning*. Berlin: Springer Verlag.

## Artificial intelligence and platform labor

Since the beginning of the 2000s, a new research perspective has opened up in digital labor. As part of the SID research team, the DiPLab (Digital Platform Labor) research program mainly focuses on the human work of data production for training, verification, and impersonation of AI solutions. A considerable section of the literature on this topic explores the nexus between workers and companies on digital platforms, in the context of North-South relationships. In the HUSH (Human Supply Chain) ANR project [5], we address the question of how European AI companies outsource their annotation and data generation activities to remote workers in African countries. The parallel TRIA (The labor of IA) project [6] deals with Central and South-American micro-workers. In a preliminary report, we highlight how European companies use a variety of outsourcing methods in order to bridge the gap between demands for "platformization" from large international groups and their clients' need for business-oriented solutions [7]. The obtained results are presented in books, conferences, journals, and content appealing to a wider audience, such as the award-winning documentary "Invisible—The Click workers" (France Télévisions) [8].

▸ [5] diplab.eu/projects-2/hush-the-human-supply-chain-behind-smart-technologies

▸ [6] diplab.eu/projects-2/tria

▸ [7] C. Belletti, D. Erdsiek, U. Laitenberger & Paola Tubaro (2021) "Crowdworking in France and Germany", *ZEW-Kurzexpertise* Nr. 21-09, Mannheim, intermediary report HUSH project, zew.de/publikationen/crowdworking-in-france-and-germany

▸ [8] france.tv/slash/invisibles/

# Coopération internationale

# International
## Cooperation

Non-exhaustive selection of our international partners

**Canada**
Institute for Quantum Computing, Waterloo
University of Ottawa, Ottawa
University of Toronto, Toronto
University of Waterloo, Waterloo

**USA**
Cisco, Mountain View, California
Columbia University, New-York
HP Labs, Palo Alto
Qualcomm, Massachusetts
Rutgers University, New Jersey
Sandia National Labs, Albuquerque
University of California (Los Angeles), Los Angeles
University of California (Santa Barbara), Santa Barbara
University of Florida, Gainesville
University of Southern California, Los Angeles

**Brazil**
Instituto Tecnológico de Aeronáutica, São José dos Campos
Universidade de Brasilia, Brasilia
Universidade de São Paulo, São Paulo
Universidade Estadual de Campinas, Campinas

# Coopération
## internationale

Une sélection non-exhaustive de nos partenaires internationaux

**Sweden**
KTH, Stockholm

**Denmark**
Aarhus University, Aarhus

**Switzerland**
EPFL, Lausanne
ETH Zurich, Zurich

**Belgium**
Université Catholique de Louvain, Louvain
Université de Liège, Liège

**Luxembourg**
Université du Luxembourg, Luxembourg

**UK**
Aston University, Birmingham
Imperial College of London, London
University College of London, London

**Germany**
Infinera Corporation, München
Max-Planck-Institut für Informatik,
Saarbrücken
Nokia Bell Labs, Stuttgart
Technische Universität Chemnitz,
Chemnitz
Technische Universität Darmstadt,
Darmstadt
Technische Universität München,
München
University of Würzburg, Würzburg

**Romania**
Universitatea Politehnica din București,
Bucharest

**Poland**
Warsaw University of Technology, Warsaw

**Russia**
ITMO University, Saint Petersburg

**Italy**
University of Trento, Trento

**Spain**
ICFO, Barcelona

**China**
Anhui University, Hefei
Chinese Academy of Science, Beijing
Shanghai Jiao Tong University, Shanghai,
ACESi
ShanghaiTech University, Shanghai
Southeast University, Nanjing
Tsinghua University, Beijing
Wuhan University, Wuhan

**Hong Kong**
Hong Kong University of Science
and Technology HKUST, Hong Kong

**Japan**
Kobe University, Kobe
NAIST, Ikoma
Tōhoku University, Sendai
Tokyo Tech, Tokyo

**India**
IIT Hyderabad, Hyderabad
Indian Institute of Science, Bangalore

**Algeria**
Université de Boumerdes, Boumerdes

**Tunisia**
Sup'Com, Tunis

**Morocco**
Université Internationale de Rabat, Rabat

**Senegal**
Cheikh Anta Diop University, Dakar

**Israel**
Bar Ilan University, Tel Aviv
Tel Aviv University, Tel Aviv

**Qatar**
Texas A&M University, Al Rayyan

**Vietnam**
Hanoi University of Science and Technology HUST, Hanoi

**Singapore**
Nanyang Technological University, Singapore
National University of Singapore, Singapore

**Australia**
University of Canberra, Canberra

**New Zealand**
University of Waikato, Waikato

# We welcomed
# **new joiners**
## in 2021

## Amal Dev Parakkat
**Associate professor**

Amal Dev Parakkat received his PhD in 2018 from the Indian Institute of Technology, Madras. He was a postdoctoral researcher at the GeoVic (STREAM) team of LIX at École Polytechnique (2019) and the Computer Graphics and Visualization (CGV) group at TU Delft (2021). Previously, he also worked as an Assistant professor at the Indian Institute of Technology, Guwahati (2020). His research interests lie in the interactive computer graphics and geometric processing domains. He is a member of the international program committee of several conferences in computer graphics, including Eurographics (S), Shape Modeling International, and Computer Graphics International.

▸ Joined the **SSH** team on September 1, 2021.

## Maria Mushtaq
**Associate professor**

Maria Mushtaq received her PhD in Information Security from the Université de Bretagne Sud (UBS) in 2019, on the subject of "Software-based Detection and Mitigation of Microarchitectural Attacks on Intel's x86 Architecture". She held a post-doctoral position at LIRMM of University of Montpellier for 2 years, with the support of a grant for excellence. Her areas of expertise lie in microarchitectural vulnerability assessment and design & development of runtime mitigation solutions against side- and covert-channel information leakage in modern computing systems. Her work focuses on cryptanalysis, constructing and validating software security components, and constructing OS-based security primitives against various hardware vulnerabilities. She is currently involved in several national and international projects.

▸ Joined the **SSH** team on September 1, 2021.

## Mathieu Fontaine
**Associate professor**

After a PhD at Inria Nancy Grand-Est on "Alpha-stable processes for signal processing" defended in 2019, Mathieu Fontaine started in October 2019 a post-doctoral fellowship at the RIKEN Center for Advanced Intelligence Project (AIP) (Wakō, Japan) and became a visiting scholar at Kyoto University. His research topics are mainly centered around audio signal processing or stochastic processes applied to signal processing and include, among others, applications in speech enhancement, speech separation, sound source localization, music separation via probabilistic heavy-tailed models and/or by deep Bayesian learning.

▸ Joined the **S2A** team on October 1, 2021.

# Ceux qui nous ont rejoint
# en 2021

### Enzo Tartaglione
**Associate professor**

Enzo Tartaglione received a joint Master's degree in Electronic Engineering (110/110 cum laude) from: Politecnico di Torino, University of Illinois at Chicago and Politecnico di Milano in 2015. He received the Alta Scuola Politecnica diploma in 2016. In 2019, he defended his PhD thesis at Politecnico di Torino, on the topic "From Statistical Physics to Deep Neural Network Algorithms". From 2019 to 2021, he held a postdoctoral position" at the University of Turin, working on the European project "DeepHealth". His research topics include neural networks pruning, compression, quantization, regularization, deep learning applied to medical image processing, data and model debiasing, privacy preservation.

> ▸ Joined the **MM** team on October 1, 2021.

### Robert Graczyk
**Associate professor**

Robert Graczyk received his MSc and PhD in Electrical Engineering from ETH Zurich in 2017 and 2021, respectively. He specializes in information theory, with a focus on information measures, inference, and control. His doctoral thesis, "Guess What?", comprises a novel set of mathematical tools that allow for the precise analysis of a broad range of variations on the classical Massey-Arikan guessing problem. His current research interests center around the fundamental limits in machine learning and control with imperfect information.

> ▸ Joined the **ComNum** team on December 27, 2021.

### Peter Brown
**Associate professor**

In May 2019, Peter Brown received his PhD in Mathematics, University of York (UK), on the topic "On constructions of quantum-secure device-independent randomness expansion protocols", where he was awarded the Anand Ramachandran Memorial Prize for the best PhD thesis in the Department of Mathematics. He joined Inria, ENS de Lyon, in December 2019, as a postdoctoral researcher. He focuses on projects that have the overarching goal of making device-independent cryptographic protocols technologically viable. This involves developing mathematical tools to analyze such protocols at a fundamental level as well as applying these tools to design secure protocols that could be implemented in an experiment.

> ▸ Joined the **IQA** team on December 13, 2021.

### Weiqiang Wen
**Associate professor**

Weiqiang Wen completed his PhD at LIP (ENS de Lyon, France) in November 2018, on the topic "Contributions to the Hardness Foundations of Lattice-Based cryptography". He then carried out his postdoctoral work in the EMSEC team of the IRISA laboratory in Rennes. His research interests span cryptography and lattice-based problems, a promising branch in post-quantum cryptography called lattice-based cryptography. His works mainly focus on understanding the security of the practical lattice-based cryptographic schemes better. This mainly involves designing better attacks on the lattice-based schemes and proving tighter hardness results for the underlying lattice problems.

> ▸ Joined the newly created **C²** team on December 15, 2021.

### Hicham Janati
**Associate professor**

Hicham Janati obtained his PhD at at Institut Polytechnique de Paris, "Advances in Optimal transport and applications to Neuroscience" in March 2021. He did a post-doc stay at CMAP, École Polytechnique (France) at Rémi Flamary's lab for six months working on domain adaptation and self-supervised learning. His research interests are optimal transport and its interactions with machine learning with applications in brain imaging and genomics. In particular, he presented Optimal transport closed form for Gaussians at the Neur IPS 2020 orals 2020, Debiased Sinkhorn barycenters at ICML 2020, and spatio-temporal optimal transport at AISTATS 2020.

> ▸ Joined the **S²A** team on December 15, 2021.

### Jean-Samuel Beuscart
**Associate professor**

Jean-Samuel Beuscart is a sociologist. After completing his PhD at ENS Cachan, he became a researcher in the social sciences department of Orange Labs, where his research focused on digital markets and digital consumption. He combines an economic sociology of markets with Internet studies in order to explore topics such as the digitization of cultural markets, the mechanisms of online cultural participation and the use of digital recommendation systems. He published "*Promouvoir les œuvres culturelles*" (2012) and "*Sociologie d'internet*" (2016) and he obtained his accreditation to supervise research (HDR) in 2019. His current work focuses on the use of digital tools to foster sustainable behaviors and on the link between digital and ecological transitions.

> ▸ Joined the **SID** team on December 27, 2021

# Among our
# young talents

## Ulrich Laitenberger
**ECOGE**

**Web**
www.ulrich-laitenberger.com

Ulrich Laitenberger's research interests lie in empirical industrial organization, competition policy and digital markets and platforms. He studies empirically potential competition problems regarding large online platforms like Amazon or Booking.com. Another one of Ulrich's research topics is how the ecosystem of online labor platforms operates and the role of intellectual property in the digital economy. To this end, he relies on the analysis of large data sets and collaborates with international researchers.

His research has been funded twice by the French National Research Agency (ANR), and recognized with prizes ("Outstanding Scientific Performance" award from the Sponsors' Association of the European Economic Research Centre (ZEM) for the article "Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection"). He is co-organizer of the Paris Conference on Digital Economics which was held in March 2022.

Since 2021, Ulrich has been a member of the Expert Group for the EU Observatory on the Online Platform Economy, who advises the Commission on main trends, analyzes potentially harmful practices and the evolution of policy measures.

### PUBLICATIONS

▸ S. Heim, K. Hüschelrath, **U. Laitenberger**, and Y. Spiegel, "The Anticompetitive Effect of Minority Share Acquisitions: Evidence from the Introduction of National Leniency Programs," *American Economic Journal: Microeconomics*, Aug. 2020.

▸ M. Hunold, K. Hüschelrath, **U. Laitenberger**, and J. Muthers, "Competition, Collusion and Spatial Sales Patterns – Theory and Evidence," *Journal of Industrial Economics*, vol. 68, no. 4, pp. 737–779, 2020.

▸ M. Hunold, R. Kesler, Reinhold and **U. Laitenberger**, "Hotel Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection" (December 2018). *ZEW* - Centre for European Economic Research Discussion Paper No. 18-059.

### CV

▸ PhD in Business Economics from KU Leuven in 2015

▸ Research associate at the Centre for European Economic Research (ZEW) in Mannheim since 2017

▸ Accredited to supervise research (HDR) from the University Pantheon-Assas in 2019

▸ Joined Télécom Paris (**ECOGE**) in May 2017

**See also**
▸ ECOGE team, page 72

## Kyriaki Niotaki
**RFM²**

Kyriaki Niotaki specializes in the design of RF, microwave and millimeter wave circuits and systems for emerging applications in the areas of Internet of Things, RFID, wireless sensing and future wireless network technologies. Her research focuses on the design of power amplifiers for wireless transceiver systems and wireless power transfer applications. She also conducts research in multi-technology energy harvesting systems to implement concepts such as the Internet of Things and smart cities.

She is the recipient of the IEEE Microwave Theory and Techniques Society (MTT-S) graduate fellowship, the recipient of the URSI young scientist award and the co-recipient of the best paper award at the 2014 IEEE RFID Technologies and Applications conference. In 2015, she was a visiting researcher at RF and Microwave Research group at University College Dublin, Dublin, Ireland. She spent a few years in industry as an RF design engineer, before returning to academia in 2018.

### PUBLICATIONS

▸ J. Mcmenamy, A. Narbudowicz, **K. Niotaki**, and I. Macaluso, "Hop-Constrained mmWave Backhaul: Maximising the Network Flow," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 596–600, May 2020.

▸ **K. Niotaki**, A. Collado, A. Georgiadis, S. Kim and M.M. Tentzeris, "Solar/Electromagnetic Energy Harvesting and Wireless Power Transmission," in *Proceedings of the IEEE*, vol. 102, no. 11, pp. 1712-1722, Nov. 2014.

▸ S. Kim, R. Vyas, J. Bito, **K. Niotaki**, A. Collado, A. Georgiadis, *et al.*, "Ambient RF Energy-Harvesting Technologies for Self-Sustainable Standalone Wireless Sensor Platforms," *Proceedings of the IEEE*, vol. 102, no. 11, pp. 1649–1666, 2014.

### CV

▸ MSc degree in Electronic Physics from Aristotle University of Thessaloniki, Thessaloniki, Greece, in 2011

▸ PhD from Polytechnic University of Catalonia (UPC), Barcelona, Spain, in 2014

▸ Visiting Researcher at RF and Microwave Research group in University College Dublin, Dublin, Ireland, in 2015

▸ RF Design Engineer, in 2015

▸ Post-Doctoral Researcher and subsequently Associate Professor/Assistant Professor at Maynooth University, Maynooth, Ireland, in 2018

▸ One semester lecturing in Maynooth International Engineering College in Fuzhou, China.

▸ Joined Télécom Paris (**RFM²**) in November 2020

**See also**
▸ RFM² team, page 38

# Parmi nos jeunes talents

## Stéphane Lathuilière
**MM**

Stéphane Lathuilière's research interests include machine learning for computer vision problems (e.g., domain adaptation, continual learning, learning with limited supervision) and deep models for image and video generation. He has published papers in the most prestigious computer vision conferences (CVPR, ICCV, ECCV, NeurIPS) and in leading journals (T-PAMI).

Stéphane was awarded a JCJC (Young Researcher) grant by the French National Research Agency (ANR), on the subject of "Online Domain Adaptation in Changing Environments" in 2020. This project will tackle issues on partial domain adaptation, structured prediction problems, new online optimization methods for fast online adaptation and continual learning methods for structured prediction tasks.

He is currently co-advising several PhD theses with industrial partners: on unsupervised domain adaptation for person re-identification for ATOS, on generative models for complex visual data for Facebook AI Research and on automatic analysis of image quality criteria on natural scenes by deep neural networks for DxOMark.

---

### 📄 PUBLICATIONS

▸ A. Siarohin, **S. Lathuilière**, S. Tulyakov, E. Ricci, and N. Sebe, "First Order Motion Model for Image Animation," in *Advances in Neural Information Processing Systems*, 2019, vol. 32.

▸ W. Menapace, **S. Lathuilière**, S. Tulyakov, A. Siarohin, and E. Ricci, "Playable Video Generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR), Jun. 2021, pp. 10061–10070.

▸ W. Menapace, **S. Lathuilière**, and E. Ricci, "Learning to Cluster Under Domain Shift," in *Computer Vision – ECCV 2020*, Cham, 2020, pp. 736–752.

▸ **S. Lathuilière**, P. Mesejo, X. Alameda-Pineda, and R. Horaud, "A Comprehensive Analysis of Deep Regression," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 9, pp. 2065–2081, 2020.

---

### 👤 CV

▸ MSc degree in applied mathematics and computer science from ENSIMAG, Grenoble Institute of Technology (Grenoble INP), France, in 2014

▸ Master's thesis in the International Research Institute MICA (Hanoi, Vietnam)

▸ PhD in mathematics and computer science in the Perception Team at Inria, obtained from the Université Grenoble Alpes (France) in 2018

▸ Post-doctoral fellow at the University of Trento in the Multimedia and Human Understanding Group, until October 2019

▸ Joined Télécom Paris (**MM**) in October 2019

**See also**
▸ MM team, page 64

## Stefano Zacchiroli
**ACES**

Stefano Zacchiroli's current research topics span digital commons, open source software engineering, computer security, and the software supply chain. Since 2021, he has been a developer of Debian, one of the most popular GNU/Linux distributions in the world. He also served as Debian project leader from 2010 to 2013 and as a Board Director of the Open Source Initiative, the international non-profit organization that determines which licenses can be considered "open source". He received the O'Reilly Open Source Award for his contributions to Free and Open Source Software in 2015.

He is co-founder and CTO of Software Heritage, an initiative launched in 2016 by Inria in collaboration with UNESCO, to build the universal archive of software source code. By collecting, preserving and sharing all available software source code, Software Heritage opens up new avenues of research in software quality improvement, software security, developer productivity, and the historical evolution of programming languages for industrial practitioners and academic researchers. These topics are at the heart of his research work at Télécom Paris.

---

### 📄 PUBLICATIONS

▸ R. Di Cosmo and **S. Zacchiroli**, "Software Heritage: Why and How to Preserve Software Source Code," in iPRES 2017 - *14th International Conference on Digital Preservation*, Kyoto, Japan, Sep. 2017, pp. 1–10.

▸ P. Boldi, A. Pietri, S. Vigna, and **S. Zacchiroli**, "Ultra-Large-Scale Repository Analysis via Graph Compression," in *2020 IEEE 27th International Conference on Software Analysis, Evolution and Reengineering* (SANER), 2020, pp. 184–194.

▸ C. Lamb and **S. Zacchiroli**, "Reproducible Builds: Increasing the Integrity of Software Supply Chains," *IEEE Software,* 2021.

---

### 👤 CV

▸ PhD in computer science from the University of Bologna, Italy, in 2007

▸ Associate Professor at University Paris Diderot in 2011

▸ Accredited to supervise research (HDR) at University Paris Diderot in 2017

▸ Joined Télécom Paris as Full Professor (**ACES**) in September 2021

**See also**
▸ ACES team, page 44
▸ software Heritage, page 86

# Editorial
## Édito

**Talel Abdessalem**
Dean of Research – Director of the LTCI

EN In 2021, Télécom Paris consolidated its very good results in the French rankings of major Engineering Schools. Thanks to Institut Polytechnique de Paris, it also strengthened its position in the international rankings of universities worldwide. These rankings naturally stem from the high quality of the teaching delivered in the various training courses of the School, the excellence of the research carried out in its laboratories and the closeness of its ties with academic and industrial partners.

This year, despite Covid-19 and its ongoing restrictions, the research teams of Télécom Paris once again surpassed themselves to ensure the continuity of their teaching activities and pursue their research projects. Despite these tough conditions, the School showed resilience and a fierce desire to maintain the vitality of its scientific activity.

The proof is that the work carried out in its two research laboratories, the LTCI (Information Processing and Communication Laboratory) and i3 (Interdisciplinary Institute of Innovation) was the subject of numerous publications and rewarded with several international prices.

*2021 also marked a strengthening of our expertise in the field of digital infrastructure security with two targeted recruitments in cryptography and the creation of the Cybersecurity and Cryptography [C²] team at LTCI, bringing together the former teams Discrete Mathematics – Coding – Cryptography (MC2) and Network Security (CCN).*

You can find out more about this new team in the following pages. You will also find a sample of projects and results from other Télécom Paris teams. With more than 150 Faculty members and more than 320 PhD students, postdoctoral researchers and research engineers, these teams contribute to pushing the limits of knowledge a little more each day in the field of information and communication sciences and technologies and their societal and economic impacts.

FR En 2021, Télécom Paris conforte sa très bonne position dans les classements français et renforce, avec l'Institut Polytechnique de Paris, sa position dans les classements internationaux. Ces résultats dans les classements découlent naturellement de la grande qualité de l'enseignement délivré dans les différentes formations de l'École, de l'excellence de la recherche menée dans ses laboratoires et de la solidité de ses relations avec ses partenaires académiques et industriels.

Cette année, placée une fois encore sous le signe de la Covid-19 et des contraintes sanitaires, les équipes de recherche de Télécom Paris se sont de nouveau dépassées pour assurer la continuité de leurs activités d'enseignement et poursuivre leurs projets de recherche. Malgré cette rude mise à l'épreuve, l'École a su faire preuve de résilience et d'une volonté farouche de maintenir le dynamisme de son activité scientifique.

Les travaux menés au sein de ses deux laboratoires de recherche, le LTCI (Laboratoire Traitement et Communication de l'Information) et i3 (Institut Interdisciplinaire de l'Innovation), ont ainsi fait l'objet de nombreuses publications et ont été récompensés de plusieurs prix internationaux.

*2021 marque également un renforcement de notre expertise dans le domaine de la sécurité des infrastructures numériques par deux recrutements ciblés en cryptographie et la création de l'équipe Cybersécurité et Cryptographie [C²] au LTCI, réunissant les anciennes équipes Mathématiques discrètes – Codage – Cryptographie (MC2) et Sécurité des réseaux (CCN).*

Vous pourrez en savoir plus sur cette nouvelle équipe dans les pages qui suivent, vous découvrirez également dans cette rétrospective un échantillon des projets et résultats issus des autres équipes de Télécom Paris, où plus de 150 enseignants-chercheurs et plus de 320 doctorants, post-doctorants et ingénieurs de recherche contribuent à repousser un peu plus chaque jour les limites des connaissances dans le domaine des sciences et technologies du numérique et de leurs impacts sociétaux et économiques.

# Strategic focuses
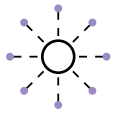
## Very Large Networks & Systems

(60-61)

**57** Permanent members

**153** Publications

**12** Thesis



## Design, Interaction, Perception

(54-55)

**14** Permanent members
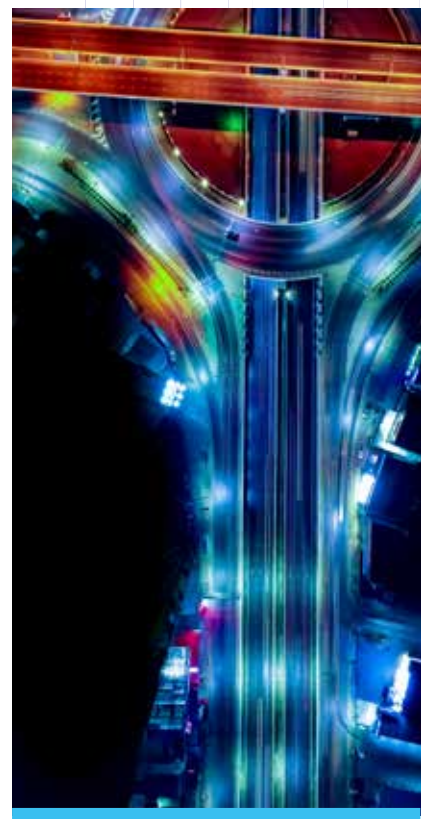
**38** Publications

**1** Thesis



## Digital Trust

(40-41)

**25** Permanent members

**62** Publications

**9** Thesis

Six strategic axes materialize the potential for transversal and interdisciplinary mobilization of Telecom Paris research, on major issues related to the digital revolution across all industrial and economic sectors.

(68-69)

# Data Science & Artificial Intelligence

**57** Permanent members

**146** Publications

**24** Thesis

(82)

**Transversal focus**

# Digital Innovation

**31** Permanent members

**99** Publications

**3** Thesis

(34-35)

**Transversal focus**

# Mathematical Modeling

## 3 main application domains

> **Content, knowledge and interactions**

> **Networks and systems**

> **Information and its ecosystems**
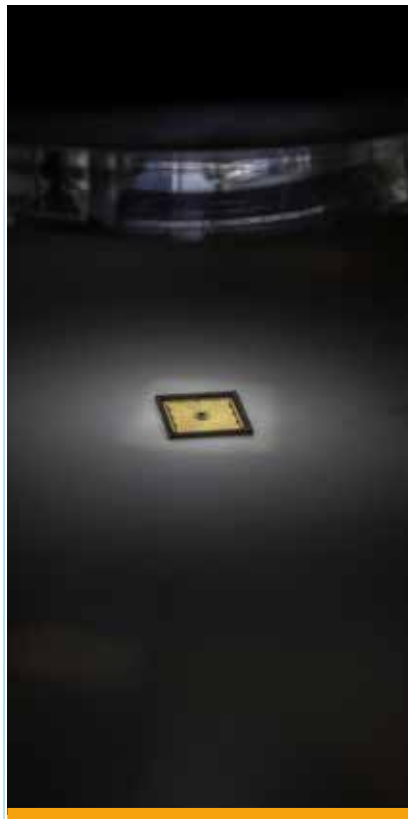
01

# Research
## Recherche

**Metrics**

Permanent members include researchers, faculty members, associate researchers, emeritus researchers and engineers. Publications include, over one year: books and book chapters, publications in international conferences with program committees and proceedings, articles in scientific journals and other publications. Thesis defenses are agreed over one year. Patents are accounted for 5 years.

The following pages present only a selection of the research currently being conducted within the various teams. Visit their websites for more information.

# LTCI

**Laboratoire Traitement et Communication de l'Information**

The LTCI brings together about 420 employees, including nearly 130 permanent research professors and more than 190 doctoral students. Organized into 15 research teams, it produces around 530 annual international publications (in scientific journals and books, or in leading international conferences). At the Institut Polytechnique de Paris, the LTCI works regularly with teams from the i3 laboratory (page 70), LIX, CMAP, Samovar and CREST. It also maintains strong links with a vast Ile-de-France ecosystem, presented on page 88 and following.

Members of the laboratory have obtained four ERC grants (since 2015), including two ongoing ERC Starting Grants: Michèle Wigger (ERC 2016; theoretical aspects of information transmission, flow management and distributed storage) and Mansoor Yousefi (ERC 2018; mathematical methods based on nonlinear Fourier theory).

The LTCI, which has been certified for the quality of its industrial research partnerships, is a member of the Carnot Télécom & Société Numérique (p. 92).
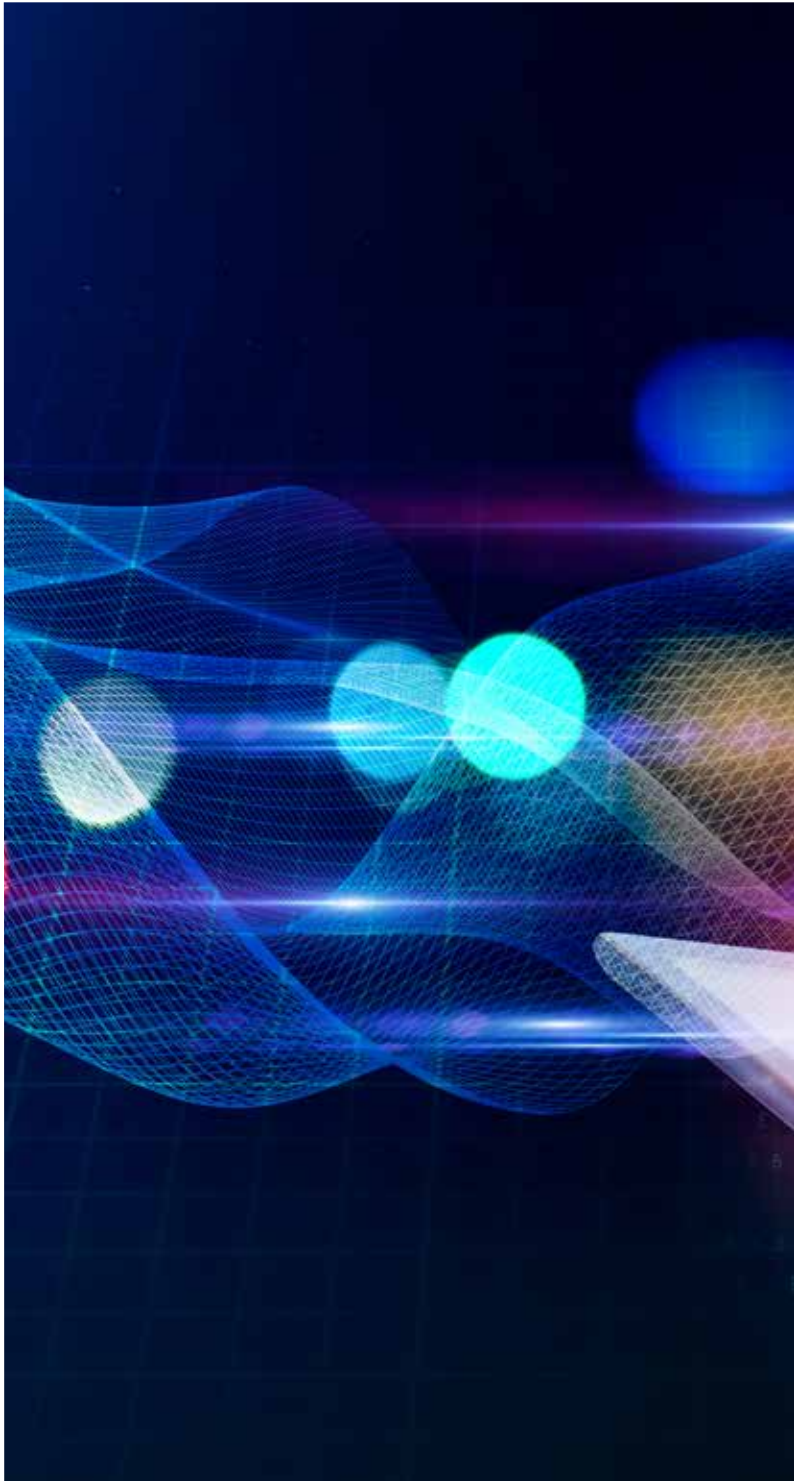
The **Information Processing and Communications Laboratory** (LTCI) is a research laboratory of Télécom Paris. It succeeds, since January 2017, the CNRS Joint Research Unit (UMR) of the same name. The LTCI was created in 1982 and is characterized by its extensive coverage of topics in the field of information and communication sciences and technology. The laboratory is also active in issues related to systems engineering and applied mathematics.

Information and communication technology becomes more firmly embedded in our society and fundamentally affects the way it operates. The objects which instantiate their highly multidisciplinary character require, in order to be developed, the understanding and exploitation of physical phenomena, the algorithmic processing of signals and data passing through them, and the way they efficiently slot into a global information system. They are socio-technical objects and as such, represent some of today's major challenges, including access to information networks and knowledge, energy efficiency and environmental constraints, privacy, security and health, as well as cooperation with artificial intelligence systems.

Research at the LTCI addresses three main domains: **Digital Communications and Electronics** – physical and hardware basis of processing and design/modeling/integration of objects in an information system; **Computer Science and Networks** – infrastructure, systems and digital networks, with a special focus on mathematical modeling, protocols and functions of next generation networks, computing architecture of systems and services, in their algorithmic and software aspects; **Images, Data and Signal Processing** – study and analysis of signals and images in all their forms, study and development of algorithms and statistical processing methods for learning, optimization and data analysis.

This cutting edge research is also carried out in industrial chairs and in several joint laboratories with our partners from academia and/or industry. The laboratory regularly organizes multidisciplinary research seminars on critical embedded systems, data science and artificial intelligence, communication systems and networks and, jointly with the i3 laboratory, on innovation issues and the socio-economic aspects of digital technologies. LTCI researchers also offer their expertise in governmental bodies, regulation authorities and in telecommunications standard-setting organizations.

# C2S

## Group leader

Patricia Desgreys

## Keywords

Smart AMS systems, Frugal signal processing, Smart radio, Cyber-physical system interfaces

## Mots-clés

Systèmes AMS intelligent, Traitement parcimonieux de signaux, Radio intelligente, Interfaces des systèmes cyberphysiques

## Web
telecom-paris.fr/c2s

## Publications
telecom-paris.fr/hal-c2s

| | |
|---|---|
| Permanent members | **5** |
| Membres permanents | |

| | |
|---|---|
| Publications | **9** |
| Publications | |

## *Historical expertise in AMS circuit design*

C2S designs Radio-Frequency communication chips and is specialized in **analog and-mixed-signal** (**AMS**) **interfaces** for 5G and IoT Systems. It concentrates specifically on the critical interface between radio communication and digital processing, combining analog electronics efficiency with digital flexibility in optimal trade-offs.

Its historical expertise in AMS circuit design spans analog-to-digital converters (ADCs), receivers and samplers. For these sub-systems, C2S have designed solutions to meet the ever-increasing demand for data rate while limiting power consumption. ASICs have been implemented in CMOS technology to demonstrate the efficiency of the solutions. C2S then broadened its theoretical approach with the development of new algorithms and digital processing dedicated to AMS systems performance. The added-value lies in specific pre- and post-processing algorithms dedicated to the linearity improvement of AMS communication interfaces.

The concept towards "smart" AMS Systems has been pushed further recently by developing **compressed sensing** (**CS**) algorithms and **cognitive radio** (**CR**) algorithms, which are useful in drastically reducing the power consumption and increasing CR security.
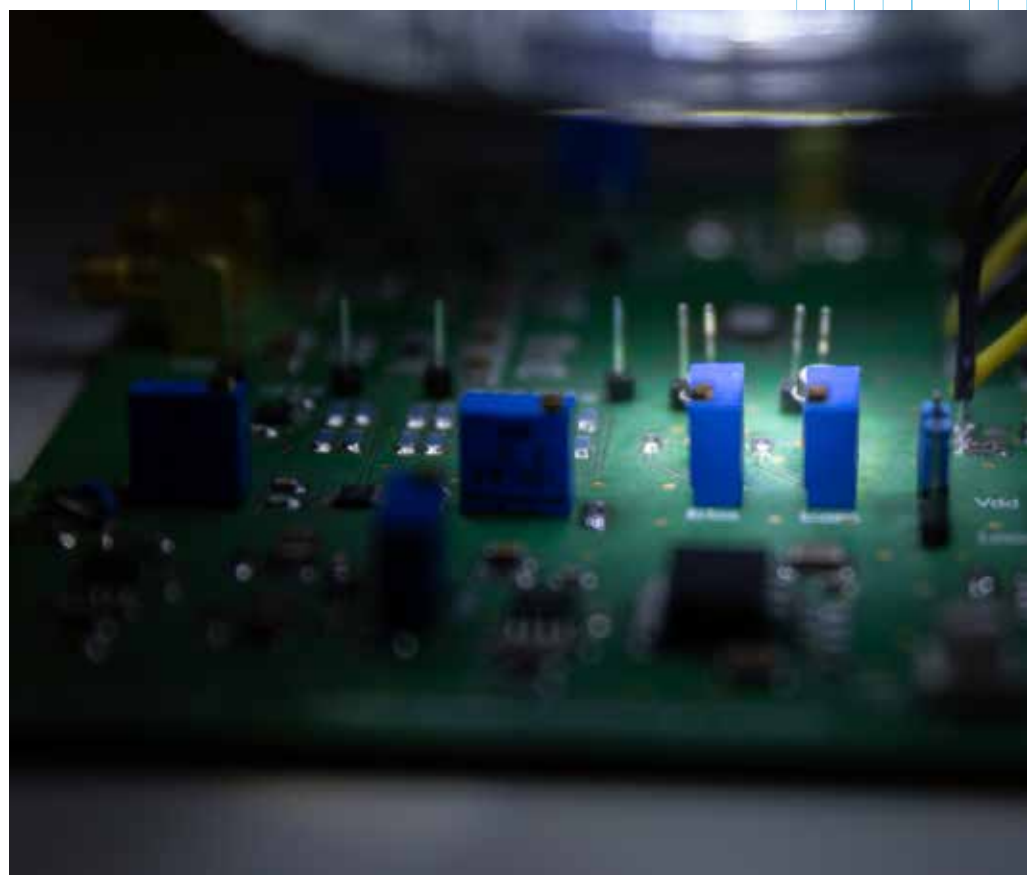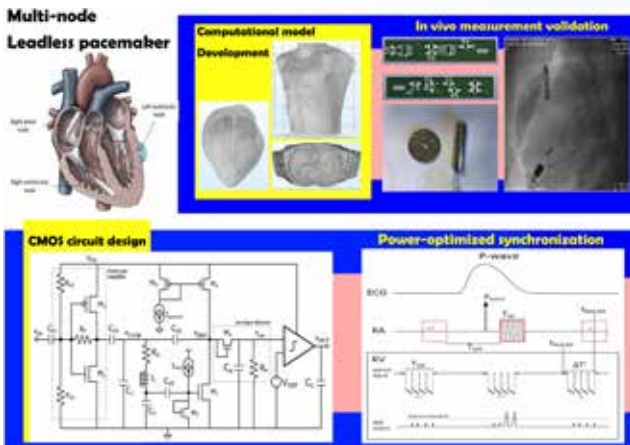
## Ultra-Low Power communication system for leadless pacemakers

Leadless cardiac pacemakers (LCPs) are small capsules directly fixed inside the heart, reducing complication risks for patients. The current LCPs can only pace a single location, thus limiting the eligible patient population. A technical solution to this limit may be to synchronize systems of LCPs. We proposed a telemetry-based solution for the synchronization of multi-node LCP systems. We used Intra-body communication (IBC), a near-field electrical communication based on electrode dipoles for its advantages in terms of size, power and cyber-security. As a first step, we characterized in article [1] the attenuation level of IBC signals propagating through the tissues. To do so, we developed a torso model for EMC simulations and verified it through in-vivo trials.

In addition, we designed an ultra-low-power, high sensitivity receiver [2]. To minimize the burden of the communication power on the longevity of the device, we also proposed a synchronization protocol that reduces the receiver duty cycle, thus achieving power-optimized synchronization for multi-node LCP systems.

▸ [1] **M. Maldari**, M. Albatat, J. Bergsland, Y. Haddab, **C. Jabbour**, and **P. Desgreys**, "Wide Frequency Characterization of Intra-Body Communication for Leadless Pacemakers," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 11, pp. 3223–3233, 2020.

▸ [2] **M. Maldari**, **C. Jabbour**, Y. Haddab, and **P. Desgreys**, "Ultra-Low Power system for atrioventricular synchronization using leadless pacemakers" URSI, The Radio Science Bulletin.

## Energy efficient information acquisition method for IoT

The Internet of Things (IoT) is currently experiencing major developments. Wireless sensor networks, which rely on wireless communications, need to be small, cheap, reliable and power efficient. Analog-to-feature (A2F) conversion is a new acquisition method designed for IoT devices to increase battery life for wireless sensors. The converter aims to extract useful features directly from the analog signal. By carefully selecting a set of features, it is possible to acquire only the relevant information required to perform classification or regression using machine learning.

Recent work has shown that the use of an A2F converter, based on the non-uniform wavelet sampling (NUWS), for electrocardiogram signals, could offer 98.5% accuracy in detecting abnormal signals. A drawback of NUWS is the large number of possibly extracted features. A power consumption model of A2F is proposed in order to perform a power-aware feature selection to optimize A2F power consumption. The figure below shows that, with the classic selection, 98.4% accuracy can be reached while requiring 20µJ (6 features), and with power-aware feature selection, 98% accuracy can be reached while requiring 3µJ (10 features). By extracting more features, the same level of accuracy can be reached, with a nearly 7-fold reduction of the energy needed for extraction.

▸ [3] **A. Back**, **P. Chollet**, **O. Fercoq**, and **P. Desgreys**, "Feature selection algorithms for flexible analog-to-feature converter," in *2020 18th IEEE International New Circuits and Systems Conference (NEWCAS)*, 2020, pp. 186–189.

▸ [4] **A. Back**, **P. Chollet**, **O. Fercoq**, and **P. Desgreys**, "Power-aware feature selection for optimized Analog-to-feature converter," *Microelectronics Journal* p. 105386, 2022.

# ComNum
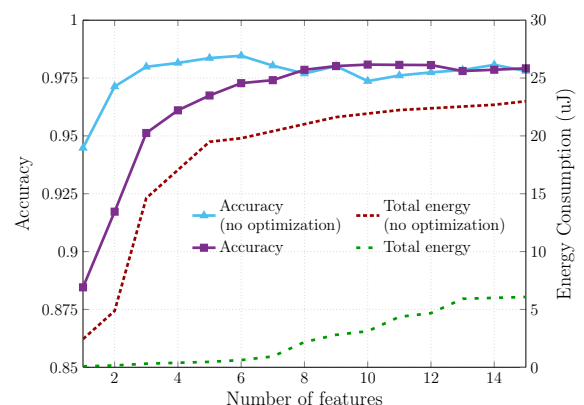
## Group leader

Ghaya Rekaya

### Keywords
Information theory, Network data processing, Coding, Distributed optimization, Security, Edge caching, Lattice theory, MIMO Coding/Decoding, Network coding, Machine learning

### Mots-clés
Théorie de l'information, Traitement de l'information sur réseau, Codage, Optimisation distribuée, Sécurité, Politique de renouvellement de cache, Réseaux euclidiens, Codage/décodage MIMO, Codage réseaux, Apprentissage machine

### See also
▶ our spinoff MIMOPT, p. 83

### Web
telecom-paris.fr/comnum

### Publications
telecom-paris.fr/hal-comnum

## Better understand and reach the fundamental limits of communication networks

| | |
|---|---|
| Permanent members / Membres permanents | **7** |
| Thesis / Thèses soutenues | **3** |
| Patents / Brevets | **18** |
| Publications / Publications | **64** |

Communication networks rely on a range of communication media that can be wireless, wired, optical etc., to transmit, retrieve and process data. Constraints are generally placed on the tasks that need to be carried out. These constraints take into account a number of performance metrics, such as complexity, reliability, energy efficiency, latency and secrecy. Depending on the nature of the network, they can be addressed in a centralized or in a decentralized way.

ComNum research efforts provide a better understanding of the fundamental limits of communication networks and means to reach these limits. Contributions span the areas of **information theory, coding, optimization, and signal processing**. The main results are obtained in the context of multi-user cellular networks (one-to-many or many-to-one communication), wireless ad hoc networks (many-to-many communication), and optical communications (one-to-one communication). For these settings, ComNum has developed insights on fundamental limits and related coding techniques for information transmission and storage over networks, on resource optimization for wireless networks, and on fundamental limits and related estimation techniques for centralized or distributed contexts.

## Local decode and update for big data compression

The storage and manipulation of huge amounts of data requires enormous resources in terms of hardware and software solutions, and ultimately energy. Hence, there is a growing need for space-optimal cloud storage solutions that allow efficient remote interactions, as frequent remote access and manipulation of large datasets generate a high volume of Internet traffic. Consider, for instance, compressing and storing DNA sequences in the cloud. If compression is handled via traditional methods, such as Lempel-Ziv, then to retrieve say a particular gene—typically a few tens of thousands of base pairs—we need to decompress the entire DNA sequence, about three billion base pairs. Similarly, the update of a small fraction of the DNA sequence requires the compressed data to be entirely updated. Motivated by this setup, [1,2] address the question "Is it possible to compress data while allowing for local retrieval and local update?" Surprisingly perhaps, the answer turns out to be a resounding "Yes." In these articles, a low complexity compression and decompression procedure is proposed which allows for maximal (entropy limit) noiseless data compression while allowing the retrieval of $k$ data symbols by querying order $k$ compressed symbols, and similarly for data update. In particular, the local retrieval or update of a single data symbol can be performed by probing a constant (independent of the size of the data set) number of compressed bits. Finally, the procedure is universal in that it operates independently of the statistics of the data symbols.

▸ [1] S. Vatedka and **A. Tchamkerten**, "Local Decode and Update for Big Data Compression," *IEEE Transactions on Information Theory,* vol. 66, no. 9, pp. 5790–5805, 2020.

▸ [2] S. Vatedka, V. Chandar, and **A. Tchamkerten**, "$O(\log \log n)$ Worst-Case Local Decoding and Update Efficiency for Data Compression," in *2020 IEEE International Symposium on Information Theory (ISIT)*, 2020, pp. 2371–2376.
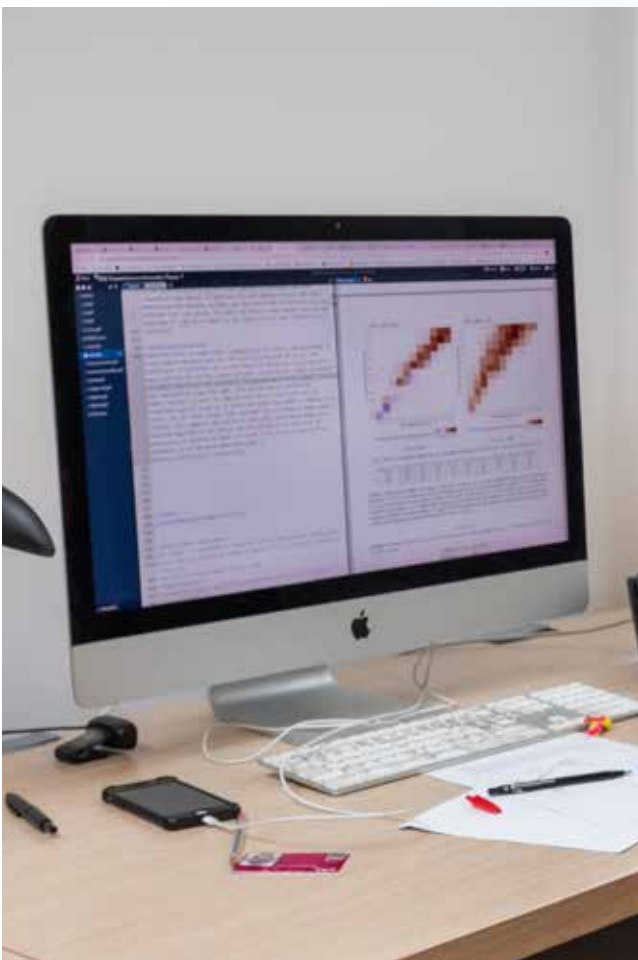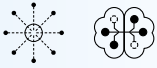
## Causal ML models: mathematical foundations and applications

*Mathematical foundations of causal machine learning (ML) models*: Causality and interpretability for machine learning (ML) models are open research problems. A causal ML model, automatically learns causal relations in the training set. In [3], we proposed a novel framework for learning Kolmogorov Models (KMs), associated with a set of binary random variables. We also derived causal relations, i.e., mathematical conditions where the realization of one variable's outcome logically implies (determines) the outcome of another. Moreover, in [4], we investigated methods that reduce computational complexity, to solve the learning problem. By leveraging strong duality and the constraints on the combinatorial learning problem, we solved its dual, using low-complexity gradient-based methods, while retaining convergence to the globally optimal solution.

*Applications of causal ML models*: In [5], we applied KMs to beam alignment for massive millimeter-wave multi-antenna systems: causal relations were used to find the best transmitter/receiver beams. In addition, due to the large predictive power of KM, a massive reduction in the signaling overhead of beam-alignment ($\approx 80\%$) without any noticeable degradation, was observed. In [3] - an extension of [6], we demonstrated the added value of the KM framework, for movie recommendation systems: the prediction performance of KMs surpassed that of all known benchmarks, and the causal relations allowed us to find groups of items for which a user liking an item implies that they like all other items in the group.

▸ [3] **H. Ghauch**, H. S. Ghadikolaei, M. Skoglund, and C. Fischione, "Learning Kolmogorov Models for Binary Random Variables," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, 2020, pp. 1204–1209.

▸ [4] Q. Duan, **H. Ghauch**, and T. Kim, "Dual Optimization for Kolmogorov Model Learning Using Enhanced Gradient Descent," accepted *IEEE Trans Signal Processing*, 2021.

▸ [5] Q. Duan, T. Kim, and **H. Ghauch**, "KM Learning for Millimeter-Wave Beam Alignment and Tracking: Predictability and Interpretability," *IEEE Access*, vol. 9, pp. 117204–117216, 2021.

▸ [6] **H. Ghauch**, H. Shokri, M. Skoglund, C. Fischione, A. Sayed, "Learning Kolmogorov Models for Binary Random Variables", *ICML Workshop on Non-Convex Optimization*, 2018.

# GTO

| | |
|---|---|
| Permanent members<br>Membres permanents | **7** |
| Thesis<br>Thèses soutenues | **2** |
| Patent<br>Brevet | **1** |
| Publications<br>Publications | **41** |

**Group leader**

Yves Jaouën

**Keywords**
Optical fiber communication, Optical networks, Nonlinear effects in optical fibers, Nonlinear Fourier transform, Advanced photonics components, Optoelectronics, Semiconductor laser dynamics, Optical fiber sensors

**Mots-clés**
Systèmes de communication optiques sur fibre, Réseaux optiques, Effets non-linéaires dans les fibres optiques, Transformée de Fourier non linéaire, Composants optiques innovants, Optoélectronique, Dynamiques des lasers à semiconducteur, Capteurs à fibres optiques.

**See also**
▸ Quantum communication infrastructure, p. 7
▸ Mid-infrared opto-electronics for high-speed data transmission, p. 61

**Web**
telecom-paris.fr/gto
**Publications**
telecom-paris.fr/hal-gto

*Bringing novel concepts in optical communication from theory to practice*

GTO conducts advanced research in high-rate fiber-optic transmission, optical network architectures, advanced lasers for communications, integrated photonics and distributed optical fiber sensors. In a context where optical communications and processing rely increasingly on digital technologies, GTO works closely with ComNum (see previous pages), as well as with world-renown international research groups in photonics and communications, placing the group in a strong position in multidisciplinary research.

Experimental demonstration of novel **concepts in optical communication** is of strategic importance. Thanks to a state-of-the-art experimental test bed, GTO have performed high-speed transmission experiments using sophisticated digital signal processors and advanced mathematical algorithms for the compensation of channel impairments.

GTO is particularly active in the following areas: high-rate communications in short-reach and long-haul transmission systems; optical networks architecture and cross-layer optimization; performance improvement in optical devices and systems for context-driven applications; quantum optics, non-linear photonics and laser dynamics.
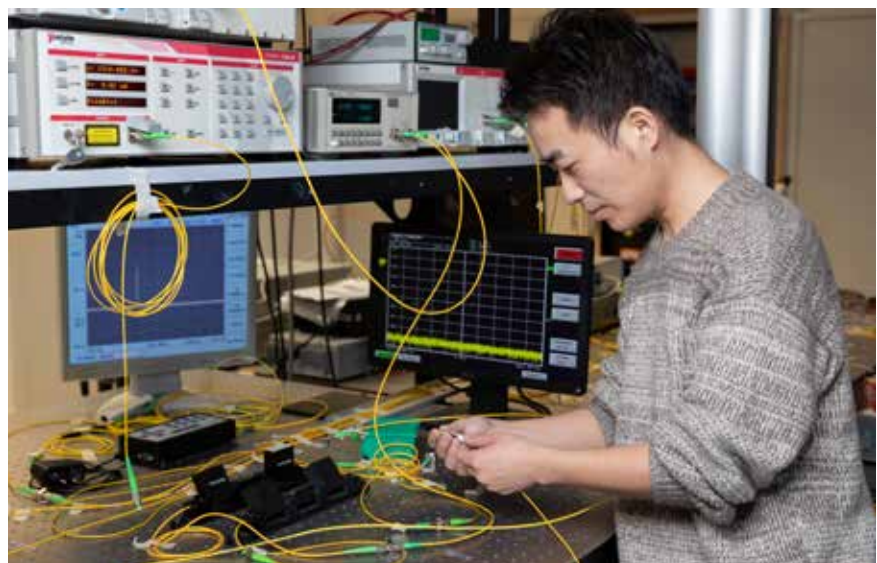
## Fiber sensors for structure health monitoring

Optical fibers can be used as a means to monitor the health of infrastructures in various areas, including civil engineering, oil and gas industries and geophysical sciences. They can also be used as a means to detect signals of interest in the surroundings. In specific contexts, exploiting already-deployed telecom fibers as distributed sensors is a better and more cost effective alternative to deploying dedicated fibers or a multitude of independent conventional sensors. Whereas the principle of distributed fiber optic vibration sensing has been suggested since 1977, optical fiber sensors have seen growing interest in recent years. This development is due to the widespread use of coherent technology for high-speed and long-distance optical transmission and the development of ultra narrow linewidth lasers for long-range phase-optical time domain reflectometry (Φ-OTDR) technics.

Distributed optical fiber sensors using Brillouin scattering have demonstrated their ability to measure changes in strain and temperature over a long distance. However, the simultaneous sensitivity of the Brillouin scattering to these two parameters is a real disadvantage. In article [1], we present a new instrument enabling independent measurement of temperature or strain using a unique fiber in a single measurement. Our method is based on the measurement of frequency, as a standard Brillouin optical time-domain reflectometry (B-OTDR) system, but also on the intensity of Brillouin scattering. As the measurement of the latter requires an accurate measurement of power levels along the fiber, it is also necessary to correct the losses with a Rayleigh OTDR measurement, which is carried out simultaneously. We present state-of-the-art results for temperature/strain separation using B-OTDR on the instrumentation of a cable into an observation well.

B-OTDR also has the advantage of requiring access to only one end of the sensing fiber, which is an important property for some applications. The coherent detection of spontaneous Brillouin scattering allows high sensitivity of temperature/strain measurements, close to one of the well-known Brillouin optical time-domain analyzer-based solutions. In-line amplification based on bidirectional EDFA modules can improve sensing distance and temperature accuracy. Article [2] demonstrates a distributed sensing over 150 km with 5-m spatial resolution and a temperature error of less than 2°C in 10 min without requiring specific BOTDR adaptation. A record performance of 0.7°C in 60 min has also been obtained, which corresponds to the best performance obtained in reflectometer configuration at this distance.

As particular cases of mechanical events, speech signals or vibrations from vehicles in the fiber cable surroundings can also be detected. In [3] and [4], we demonstrate the detection of speech signals, footsteps and vehicles through a novel polarization-multiplexed coded differential-phase OTDR interrogation of standard telecom fibers in deployed cables. This is done by measuring the induced phase variation of the optical probe signal with respect to the sound pressure level of the perturbation. These experiments pave the way for numerous applications of distributed acoustic sensing where a high sensitivity is required.

▸ [1] P. Clément, **R. Gabet**, V. Lanticq, and **Y. Jaouën**, "B-OTDR Solution for Independent Temperature and Strain Measurement in a Single Acquisition," *Journal of Lightwave Technology*, vol. 39, no. 18, pp. 6013–6020, 2021.

▸ [2] P. Clément, **R. Gabet**, V. Lanticq, and **Y. Jaouën**, "Enhancement of Sensing Range of Brillouin Optical Time-domain Reflectometry System up to 150 km with In-Line Bi-Directional Erbium-Doped Fibre Amplifications," *Electronics Letters*, vol. 57, no. 3, pp. 142–144, 2021.

▸ [3] W. Tomboza, S. Guerrier, **E. Awwad**, and C. Dorize, "High Sensitivity Differential Phase OTDR for Acoustic Signals Detection," *IEEE Photonics Technology Letters*, vol. 33, no. 13, pp. 645–648, 2021.

▸ [4] C. Dorize, **S. Guerrier**, **E. Awwad**, **P. Nwakamma** *et al.*, "An OFDM-MIMO Distributed Acoustic Sensing over Deployed Telecom Fibers," *proceedings of Optical Fiber Communications* (2021), paper W7C.2.

## Review of quantum nanostructures for information and communication technologies

Quantum nanostructures with low dimensionality like quantum dots and quantum dashes are semiconductor atoms that enable the realization of high-performance photonic devices. The following article reviews recent findings and prospects on nanostructure based light emitters, in order to allow scientists, researchers and engineers to make informed judgments about their use, for applications ranging from embedded silicon technologies to future quantum information systems. It focuses on photonic devices grown on native substrates (InP and GaAs) as well as those that are grown heterogeneously and epitaxially on silicon substrate. This work has been carried out by a research team under the direction of Frédéric Grillot.

▸ **F. Grillot**, **J. Duan**, B. Dong, and **H. Huang**, "Uncovering recent progress in nanostructured light-emitters for information and communication technologies," *Light: Science & Applications*, vol. 10, no. 1, p. 156, Jul. 2021.

LTCI is an active member of the LMH (Labex Mathématique Hadamard), which covers the full spectrum of research in mathematics, from pure and fundamental studies to applied mathematics.

Part of LTCI PhD students are registered in the doctoral school "Mathematics-Hadamard", a shared Doctoral School between Université Paris-Saclay and Institut Polytechnique de Paris.

## Generative networks for image restoration, enhancement and editing
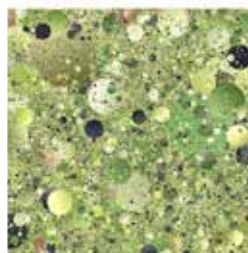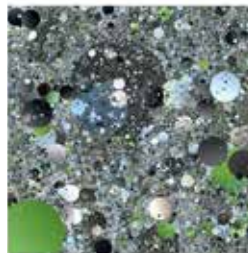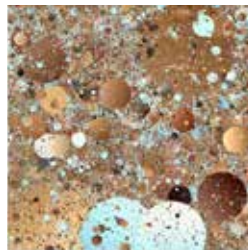


A key ingredient of image restoration methods is an a priori hypothesis about image regularity. Recently, deep neural networks have achieved impressive results in all fields of image restoration: denoising, single image super-resolution, deconvolution, etc. However, such networks usually require large databases and need to be retrained for each new modality or specific imaging device. Article [1] shows that these models can be efficiently trained from synthetic image databases, based on a mathematical model grounded in physical priors and depending on few parameters. We show that an occlusion-based dead leaves model (which are traditionally used for evaluating the preservation of textures), equipped with a scaling size distribution, is sufficient to reach near state-of-the-art restoration performances, for both tasks of denoising and single image super-resolution. We also show that this model can be efficiently combined with natural image databases to enhance the capacity of deep neural networks to preserve details, without impairing their classic performance evaluation.

This work offers interesting perspectives. Firstly, the mere structure of deep neural networks is adapted to image restoration tasks and they can be made near-optimal on the basis of just a few principles and hyper-parameters. Furthermore, the learning database we used has the potential to be modified according to specific acquisition devices and in particular to their point spread function, dynamic range, noise modality, etc. This opens the way to flexible, generic and relatively light learning schemes.

▶ [1] **R. Achddou**, **Y. Gousseau**, and **S. Ladjal**, "Synthetic Images as a Regularity Prior for Image Restoration Neural Networks," presented at the *Eighth International Conference on Scale Space and Variational Methods in Computer Vision* (SSVM), Cabourg (virtuel), France, May 2021.

# Modeling:
## a transdisciplinary tool
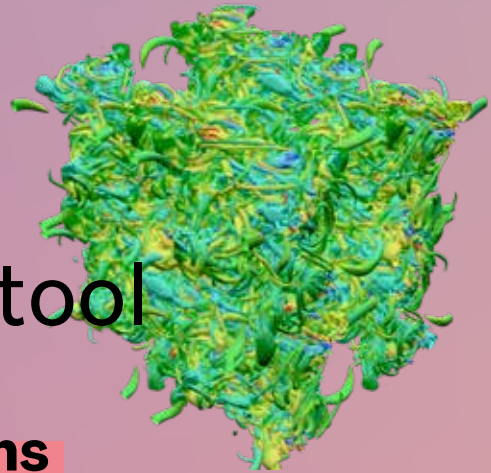


## Mathematical modeling aims to understand and facilitate simplification, optimization, representation and visualization

### Non-smooth stochastic optimization

Stochastic approximation algorithms that operate on non-convex and non-smooth functions have attracted a great deal of attention, owing to their numerous applications both in machine learning and in high-dimensional statistics. The archetype of such algorithms is the so-called stochastic subgradient descent (SGD). In [2], we establish the non-convergence of the SGD to the critical points called active strict saddles by Davis and Drusvyatskiy in a recent paper where they tackle the problem of the spurious points (local maximizers or saddle points) avoidance. Active strict saddles lie on a manifold $M$ where the function $f$ has a direction of second-order negative curvature. Off this manifold, the norm of the Clarke subdifferential of $f$ is lower-bounded. We require two conditions on $f$. The first assumption is a Verdier stratification condition, which is a refinement of the popular Whitney stratification. The second assumption, termed *the angle condition*, allows the distance of the iterates to $M$ to be controlled. When $f$ is weakly convex, our assumptions are generic. Consequently, with the help of the Verdier and the angle conditions, and generically in the class of definable weakly convex functions, the SGD converges to a local minimizer.

In the process, our first assumption allows us to establish a reinforced version of the so-called projection formula of Bolte *et al.* for Whitney stratifiable functions, which is of independent interest. We refine the projection formula to the case of definable, locally Lipschitz continuous functions by establishing a Lipschitz-like condition on the (Riemannian) gradients of two adjacent strata.

▸ [2] **P. Bianchi**, W. Hachem, and S. Schechtman. "Stochastic Subgradient Descent Escapes Active Strict Saddles," preprint.

### Topology and geometry of cellular networks

Exploring and understanding complex structures such as random graphs is a difficult and rich problem that has motivated a large literature in recent years. The natural first step when faced with large graphs is to look for cluster or community structures, i.e., a partition where the connectivity within a class is greater than the connectivity between classes. There are several ways to do this, such as modularity clustering or spectral clustering. The theory behind the latter method is particularly popular because it relates the topology of the graph to random nearest neighbor walks on it.

In her PhD thesis defended on December 10, 2020 [3], Zhihan Zhang defines first a new Markov chain on simplicial complexes and establishes some key properties over it. Then, some applications of the random walk are explored, namely random walk based hole detection and simplicial complex kernels. In both cases, novel algorithms are proposed, and simulation results show their efficiency.

▸ [3] **Zhihan Zhang**, "Random Walk on Simplicial Complexes," PhD thesis, 2020.

# LabSoC
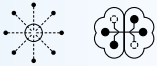
**Group leader**

Ludovic Apvrille

**Keywords**
Architecture exploration, Formal validation, Model engineering, Safety, security and performance, Security of electronic circuits, Software/hardware architecture for signal processing applications

**Mots-clés**
Exploration d'architecture, Vérification formelle, Ingénierie des modèles, Sûreté de fonctionnement, sécurité des systèmes informatiques, Sécurité des circuits électronique, Architecture logicielle / matérielle pour les applications de traitement du signal
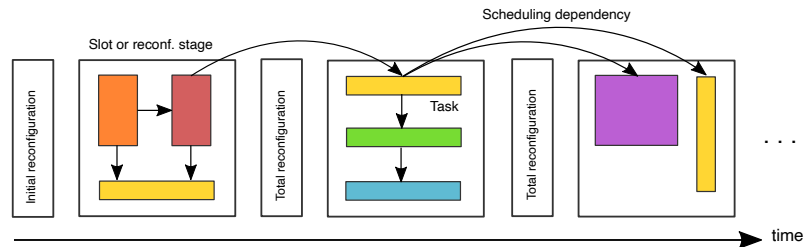
**Web**
telecom-paris.fr/labsoc
**Publications**
telecom-paris.fr/hal-labsoc

*Definition of methods, techniques, languages and tools for complex integrated and embedded systems*



LabSoC research activities focus on design, modeling and validation. The lab also analyzes and proposes new architectures for embedded systems. The main objective is to develop models and methods for better formal verification and more automatic programming of safe and secure embedded systems.

More precisely, LabSoC focuses on several research topics. Firstly, new methods, techniques, languages and tools for complex integrated and embedded systems are developed. This involves abstraction, formal methods, very fast simulation, design space exploration, handling of both safety and security aspects, and code generation from abstract models. Secondly, work is conducted on hardware security, and paranoid Systems on Chip with zero trust in their hardware and software environment. Emphasis is also placed on the confidentiality and the integrity of external communications and storage.

A last research topic concerns the digital signal processors for the software-defined radio. This involves the study of the trade-off between flexibility and energy efficiency, the associated software design environments, the abstract modeling of waveforms and hardware architecture, and the automatic generation of control software.

LabSoC contributions, based on **five platforms** (TTool, Alligator, Embb, SecBus, Rover) have first been defined internally before being applied in academic, collaborative projects (FUI Netcom, H2020 AQUAS) and industrial collaborations (Freescale, VEDECOM, Nokia, Engie). LabSoC is hosted by EURECOM at Sophia-Antipolis.

| Permanent members<br>Membres permanents | **5** | Thesis<br>Thèse soutenue | **1** | Publications<br>Publications | **10** |
|---|---|---|---|---|---|

## State space reduction for deadlock-free scheduling of signal processing applications in embedded systems

In many multi-processor systems-on-chip (MPSoCs) architectures for data processing, memories are limited in size and are shared. Scheduling of tasks must avoid memory shortage, otherwise deadlocks may occur. To scale with the growing complexity of systems—modern systems such as 4G and 5G require dynamic adjustments (modulation, bandwidth, channel-coding rate) that advocate for dynamic scheduling—existing analysis approaches typically renounce to explore all possible schedules. Consequently, run-time environments execute only a limited number of schedules that can be easily guaranteed deadlock-free.

In the PhD manuscript [1], we propose a new technique to reduce the state space for deadlock analysis. To reduce the complexity of deadlock analysis, we represent scheduling state spaces with labeled transition systems (LTSs) and we prove a fundamental theorem. Thanks to this, we show that analyzing a small subspace is sufficient to identify all state transitions that inevitably lead to deadlocks. Our evaluation reveals that the size of the systems we can study increases by more than three orders of magnitude (in terms of the number of scheduling states to analyze) compared to traditional techniques. Our work is applicable to systems with any number of memories.

▸ [1] **Benjamin Dauphin**, *"Analyses de vivacité et environnement logiciel intelligent pour la gestion mémoire d'applications flux de données"* ("Liveness Analysis Techniques and Run-Time Environment for Memory Management of Dataflow Applications"), PhD thesis from *Institut Polytechnique de Paris,* held on March 16, 2021.
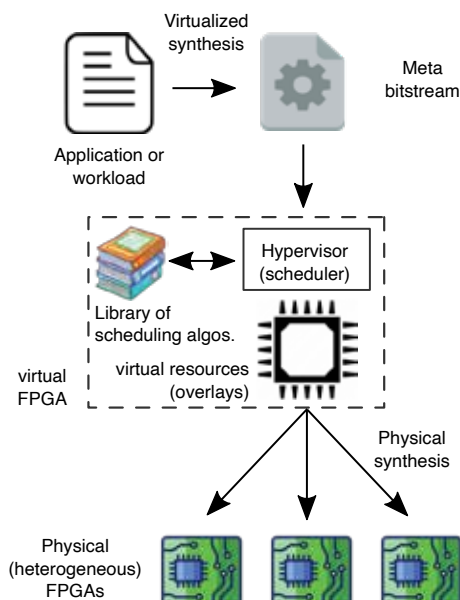
## Efficient scheduling of applications on reconfigurable hardware distributed in compute nodes

In modern cloud data centers, reconfigurable devices can be directly connected to a data center's network. For instance, field programmable gate arrays (FPGAs) are rented to users to accelerate the execution of data-intensive workloads, e.g., scientific computing, financial analysis, video processing, machine learning. Novel scheduling solutions are needed to maximize the utilization (profitability) of FPGAs, e.g., reduce latency and resource fragmentation.

Algorithms that schedule groups of tasks (clusters, packs), rather than individual tasks (list scheduling), suit the way FPGAs operate. Here, groups of tasks that execute together are interposed by hardware reconfigurations. It has been demonstrated that the classic resource constrained scheduling problem is a strongly NP-hard problem. Thus, much attention has been dedicated by the FPGA community to the design of heuristics.

In article [2], we propose an efficient heuristic based on a novel method for grouping tasks. These are gathered around a high-latency task that hides the latency of remaining tasks within the same group. The formulation of our heuristic is generic and considers $k$-dimensional models of resources whose requests do not depend on scheduling choices and are constant in time.

We evaluated our solution on a benchmark of almost 30,000 random workloads, synthesized from realistic designs (i.e., topology, resource occupancy). We compared our contribution to the Next-Fit version of the heterogeneous earliest finish time heuristic (HEFT-NF), that is the best variant of HEFT for FPGAs. For this test bench, on average, our heuristic produces optimum makespan solutions in 71.3% of the cases. It produces solutions for moderately constrained systems (i.e., the deadline falls within 10% of the optimum makespan) in 88.1% of the cases.

▸ [2] **M. Bertolino**, **R. Pacalet**, **L. Apvrille**, and A. Enrici, "Multi-Resource Scheduling for FPGA systems," *Microprocessors and Microsystems: Embedded Hardware Design (MICRO)*, vol. 87, p. 104373, Nov. 2021.

# RFM²
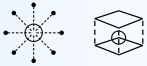
**Group leader**

Jean-Christophe Cousin

**Keywords**
Theory and modeling of RF systems, Antenna and RF circuit design, RF instrumentation, Metamaterials and metasurfaces

**Mots-clés**
Théorie et modélisation des phénomènes et des systèmes RF, Conception des antennes et des circuits RF, Instrumentation RF, Metamatériaux et metasurfaces

**Web**
telecom-paris.fr/rfm2

**Publications**
telecom-paris.fr/hal-rfm

## *Addressing long term needs independent from short term technical inflation and hype*

RFM² conducts research to connect humans or machines with wireless systems, whether mobile or embedded, with sensing mechanism to adapt to their environment applications. It addresses long term needs such as wireless communications independent from short term technical inflation and hype, national security and public health. It concentrates expertise on radio front ends with complementary skills that cover circuits and antenna designs, MMIC and antenna characterization, wave propagation and wave human interactions model up to 110 GHz.

Research is organized along three main axes: **energy-efficient RF front-end systems, smart and small antenna design as well as metamaterials and metasurfaces design, radio channel modeling and localization**. This covers the design of components and subsystems for radio links involved in 5G, IoT, satellites, localization, reliable and secured massive communications. The performance of these components and systems is evidenced by the innovative characterization methods in microwaves and millimeter wave domains. RFM² also aims to provide optimal connectivity with respect to cost, energy efficiency, performance and flexibility, in any environment.

Work also focuses on the major challenge of reducing exposure to below the legal limits. One of the RFM² strengths is to formalize both the behavior of components and subsystems and to determine the interaction between waves and humans. In this respect, RFM² is a co-founder of the internationally recognized C2M Chair (Characterization, Modeling and Control of Exposures), a multidisciplinary research framework that is both technological—numerical computation, statistics, antennae, network architecture—sociological and philosophical. This contributes to the characterization, the analysis and the model of exposure to electromagnetic waves induced by telecommunication systems and networks, as well as the study of how the risk linked to this exposure is perceived.
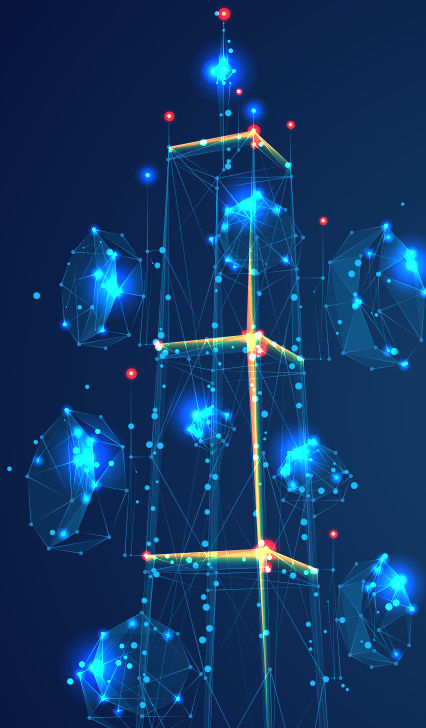
Permanent members
Membres permanents **8**

Thesis
Thèse soutenue **1**

Publications
Publications **10**

## 2D localization by multistatic UWB radars

Indoor localization has gained interest due to the vast range of services it can offer. Many positioning technologies can be used in an indoor localization system depending on the intended accuracy, precision, cost, etc. As the global navigation satellite system is not accurate enough to provide specific services, complementary systems have been developed to ensure reliability and accuracy, even in difficult conditions (multipaths, shadowing, etc.). One solution consists in using the time propagation of impulse radio-ultra wideband (IR-UWB) signal by exploiting wide frequency bandwidths and naturally increasing the accuracy.

Over the years, numerous methods based on time of arrival estimation of the UWB signal have been developed to calculate the exact range between a tag and a sensor. However, practical limitations of these algorithms prevent the localization system from achieving the desired high resolution. At the same time, a trade-off between the required accuracy, resolution, and complexity must be taken into account. We propose in article [1] an alternative solution of low complexity based on a phase correlation method adapted to IR-UWB signals, improving the accuracy and resolution of the azimuth angle measurement for an interferometric localization system. The main advantage of such an interferometric method is that the position of a person/object in a room without obstacles can be obtained using only one fixed anchor instead of a well-known constellation.

▶ [1] **N. Awarkeh**, **J.-C. Cousin**, M. Muller, and N. Samama, "Improvement of the Angle of Arrival Measurement Accuracy for Indoor UWB Localization," *Journal of Sensors*, vol. 2020, p. 2603861:1-2603861:8, Jun. 2020.
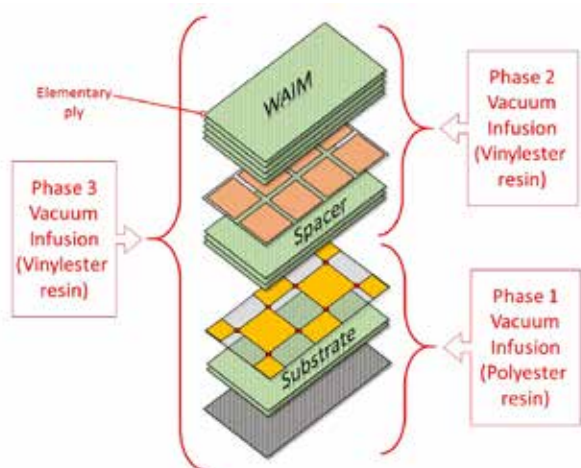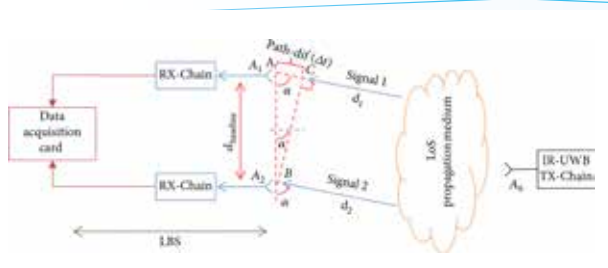
▶ Phases of manufacture of composite dielectric panel by the industrial process of molding by vacuum infusion.

## Self-complementary surface with low signature for naval applications

The SAFASNAV (*Surfaces Auto complémentaires à Faible Signature pour applications NAVales*) project developed by Naval Group in partnership with ONERA, Télécom Paris and the SME SART, is an industrial research project started in April 2017. It aims to design and incorporate metasurfaces (research theme 1) in civil and/or military ships, or wind turbines, built from composite materials, by carefully selecting the components of the structural plies and the insertion processes (research theme 2), in order to reduce the reflectivity of radar waves.

The balance between electromagnetic and mechanical performance is a constant concern, as is industrial feasibility. The concepts of metasurfaces were optimized to allow adaptation to the environment and to the requirements of the selected application case: performance required, available materials, integration constraints.

Using iterative simulations, new materials and processes were evaluated. A database was developed, including structural composite materials with low dielectric permittivity, a carbon fiber ground plane, and processes allowing the control of permittivity and thickness of structural layers. Different processes of implementation and integration of periodic patterns in walls have been evaluated and optimized. The role of the substrates used for the transfer was also taken into account as this can play a determining role in the manufacture and the performance of the absorbing walls. Nine demonstrators of absorbing panels have been fabricated and evaluated on different electromagnetic characterization benches (monostatic and bistatic reflectivity measurements).

The SAFASNAV project showed promising results. It was thus possible to build an absorbing structural wall by incorporating two metasurfaces in the external skin of a sandwich structure. The level of performance, both in terms of absorption level and bandwidth, was significant, as was robustness in incidence. The contribution of innovative technologies such as silkscreen printing has also been highlighted in the implementation of coatings with very low signatures. In addition, being able to choose between different technologies offers a considerable advantage in terms of security of supply.

First results relative to the concept have been already published [2] and the proof of the concept will be published in 2022.

▶ [2] **O. Rance**, **A.-C. Lepage**, **X. Begaud**, M. Soiron, A. Barka, and P. Parneix, "Analysis and Optimization of a Wideband Metamaterial Absorber Made of Composite Materials," *Applied Physics A*, vol. 125, no. 5, p. 358, Apr. 2019.

## TrustShare: a Blockchain-oriented innovation Chair

Blockchain can be viewed as a mechanism to implement trustworthy access to shared data in systems with mutual distrust. System participants may have conflicting interests and might even be willing to cheat, but the implementation is expected to ensure that they share data and exchange assets in a consistent, available and fair way. As originally conceived, blockchains solve a very hard problem in a very hard model. Not surprisingly, existing solutions are notoriously slow and inefficient. The goal of TrustShare Chair, initiated in 2021, is to discover novel efficient data-sharing implementations, by considering weaker, though practical variants of the problem and/or stronger, though realistic variants of the model. The Chair sponsored by Mazars and CDC (Caisse des Dépôts), is headed by Petr Kuznetsov (**ACES**).

▸ https://trustshare.telecom-paris.fr/

## Algorithms for distributed and concurrent systems

The articles refered below address the theory of computation, specifically in distributed computing. Here, consensus, by far the most studied problem, allows multiple processes to unambiguously agree on a single value. It allows a group of processes to act as a single correct process, despite crashes or even malicious behavior of some of them. In this context, practical Byzantine fault tolerance, the most widely known Byzantine fault-tolerant consensus protocol, is known to take three message delays to decide a value. We show in article [1] a new result: the optimal two-step latency tolerating $f$ Byzantine failures can be obtained with $5f - 1$ processes, which is a tight lower bound.
Consensus is not always a panacea. Indeed, it is not always necessary for implementing asset transfer, where more efficient, asynchronous solutions can be built using reliable broadcast (quorum systems) instead of consensus. In article [2], we thus present a permissionless and asynchronous asset-transfer implementation, based on a weighted proof-of-stake mechanism. Our protocol provides its guarantees in asynchronous environments where less than one third of the total stake is owned by malicious parties.

▸ [1] **P. Kuznetsov**, A. Tonkikh, and Y. X. Zhang, "Revisiting Optimal Resilience of Fast Byzantine Consensus," in *Proceedings of the 2021 ACM Symposium on Principles of Distributed Computing*, New York, NY, USA, 2021, pp. 343–353.

▸ [2] **P. Kuznetsov**, Y.-A. Pignolet, P. Ponomarev, and A. Tonkikh, "Permissionless and Asynchronous Asset Transfer," in *35ᵗʰ International Symposium on Distributed Computing, DISC 2021, October 4-8, 2021, Freiburg, Germany (Virtual Conference)*, 2021, vol. 209, p. 28:1-28:19.

## Complex autonomic systems

The architecture of coordination mechanisms is central to the performance and behavior of (self-)integrated systems across natural, socio-technical and cyber-physical domains. Multi-scale coordination schemes are prevalent in large-scale systems with bounded performance requirements and limited resource constraints. However, theories to formalize how coordination can be implemented across multi-scale systems are often domain-specific, lacking generic, reusable principles. In these systems, feedback among system entities is a key component to coordination.

We introduced in previous work a feedback-centered definition of multi-scale systems through the multi-scale abstraction feedbacks (MSAF) design pattern. In article [3], we refine it, describing a feedback cycle across scales as one where information flows bottom-up and top-down through five actions: state information communication, state information abstraction, information processing, control information communication, and adaptation from control information. We focus on analyzing coordination mechanisms with exogenous macro-entities, but are building towards a general theory of feedback in multi-scale system. In article [4], we go further and expand on MSAF to include timing concerns and illustrate their effects via two models, both generic and applicable to various domains: a hierarchical oscillator and a hierarchical cellular automaton. Our contribution sets a basis for developing a comprehensive theory of timing in multi-scale feedback systems, while also providing useful insights for designing and controlling such systems.

Article [5] proposes a fair self-adaptive and socially acceptable clustering system that compromises between the communication costs of a cellular network and the performance of a vehicle-to-vehicle (V2V) network, at runtime, depending on traffic conditions. Indeed, in such systems, only some members can directly access the cellular network, while other members communicate with these special elected members using IEEE 802.11p. Our approach is based on the theory of distributed justice to ensure a fair distribution over time of the responsibilities of being elected member and hence a better social acceptability.

▸ [3] **A. Diaconescu**, L. J. D. Felice, and P. Mellodge, "Exogenous Coordination in Multi-Scale Systems: How Information Flows and Timing Affect System Properties," *Future Generation Computer Systems*, vol. 114, pp. 403–426, 2021.

▸ [4] P. Mellodge, **A. Diaconescu**, and L. Di Felice, "Timing Configurations Affect the Macro-properties of Multi-scale Feedback Systems," presented at the 2nd IEEE Intl. *Conf. on Autonomic and Self-Organising Computing* (ACSOS), Washington (DC), USA (virtual), Oct. 2021.

▸ [5] J. Garbiso, **A. Diaconescu**, **M. Coupechoux**, and B. Leroy, "Fair Self-Adaptive Clustering for Hybrid Cellular-Vehicular Networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 2, pp. 1225–1236, Feb. 2021.

**See also**
▸ All the Chairs presented in page 10 et 11 are related to the Digital Trust issues. The three Chairs on page 11 fit into the activities of the "Operational AI Ethics" program, a multidisciplinary project which is structured around five research themes: Bias and fairness, Explainability, Responsibility, Governance and regulation, AI and public benefit.

# Smart and Trustworthy
## digital infrastructures

**Safety, security and responsibility are signs of trust**

## Internet Of Secure Elements

Article [6] introduces the Internet of secure elements (IOSE) infrastructure. IOSE is a set of secure elements providing TLS servers, communication interfaces and is identified by its name. It enables the deployment of a personal hardware secure module based on a secure element (SE) grid, providing secure storage and tamper-resistant computing resources with evaluation assurance levels ranging from EAL5+ to EAL6+. Storage capacity is limited to about 100KB, and computing time is in the range of a few hundred milliseconds. IOSE resources are protected by TLS-SE servers working with pre-shared-keys. The service plane has a pure TLS network interface. The administration plan uses remote APDU call secure servers in order to download secure element applications, identified by TLS server name. Secure element technology and components, the state of art for secure element servers and application deployment model are discussed.

▸ [6] **P. Urien**, "Towards Internet of Secure Elements," accepted at *IEEE Consumer Communications & Networking Conference 2022*, see also https://datatracker.ietf.org/doc/draft-urien-coinrg-iose

## No autonomous vehicles without trust

Trusting autonomous vehicles needs an interdisciplinary approach. Cybersecurity, risk analysis, cryptography and agility, identity management, authentication, and protection of data, data flow and personal data are thus investigated within the Connected Cars and Cyber Security Chair (page 10). Appropriation and uses, sources of trust and mistrust, sources of comfort and discomfort are studied thanks to the guided imaginary projection methodology (see Mobility, autonomous vehicles, page 75).

Diversity and inclusion must also be taken into consideration. In 2020, Jan Gugenheimer (DIVA team) explored an inclusive user-centered design for vehicle-pedestrian communication (VPC), beneficial for both vision impaired and seeing pedestrians. The VPC between groups of abilities were both studied on the ground and in virtual reality. It was shown that trust and understanding is enhanced and cognitive load reduced when all relevant vehicles communicate.

▸ [7] M. Colley, M. Walch, **J. Gugenheimer**, A. Askari, and E. Rukzio, "Towards Inclusive External Communication of Autonomous Vehicles for Pedestrians with Vision Impairments," in Proceedings of *the 2020 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 2020, pp. 1–14.

# SSH

**Group leader**

Jean-Luc Danger

**Keywords**
Architecture and design methods for embedded systems, Hardware security of embedded systems against physical and cyber attacks, Hardware reliability of embedded systems

**Mots-clés**
Architectures et méthodes de conception de systèmes embarqués, Sécurité matérielle des systèmes embarqués contre les attaques physiques et cyber, Fiabilité matérielle des systèmes embarqués

**See also**

**Web**
telecom-paris.fr/ssh
**Publications**
telecom-paris.fr/hal-ssh

## From mathematics to technology

| | |
|---|---|
| Permanent members / Membres permanents | **10** |
| Theses / Thèses soutenues | **6** |
| Patents / Brevets | **3** |
| Publications / Publications | **29** |

SSH conducts research on architectures and methods to **design efficient embedded systems and digital electronic circuits** to meet the requirements emerging from the IoT, 5G, autonomous car, smart building, etc. In addition to traditional properties of complexity, energy efficiency, operational speed, low cost and flexibility, SSH concentrates on **Security** and **Safety** constraints—cyber and physical attacks, failures—to propose robust architectures—processors, cryptoprocessors, accelerators, randomness generation—and analysis tools. These properties are becoming vital to tackle the numerous threats linked to the rising number of applications requiring a high level of trust and autonomy in hostile environments.

The study of architecture is particularly helpful when it is associated with solid validation methods. This is why many studies are related to the security/safety analysis which can take many forms: abstract (formal), machine learning, simulation and real circuit such as FPGAs and custom ASICs fabricated using CMOS technologies.

The research of SSH lies at the crossroads of industrial requirements—future technologies but also significant improvement of existing products—and new digital technologies—STT RAM, FD-SOI—with strong international collaborations. SSH scientific output consists in publications and theses, some of which can be directly applied in industry. This is the case with autonomous cars, for instance, through the C3S "Connected Cars & Cybersecurity" Chair.
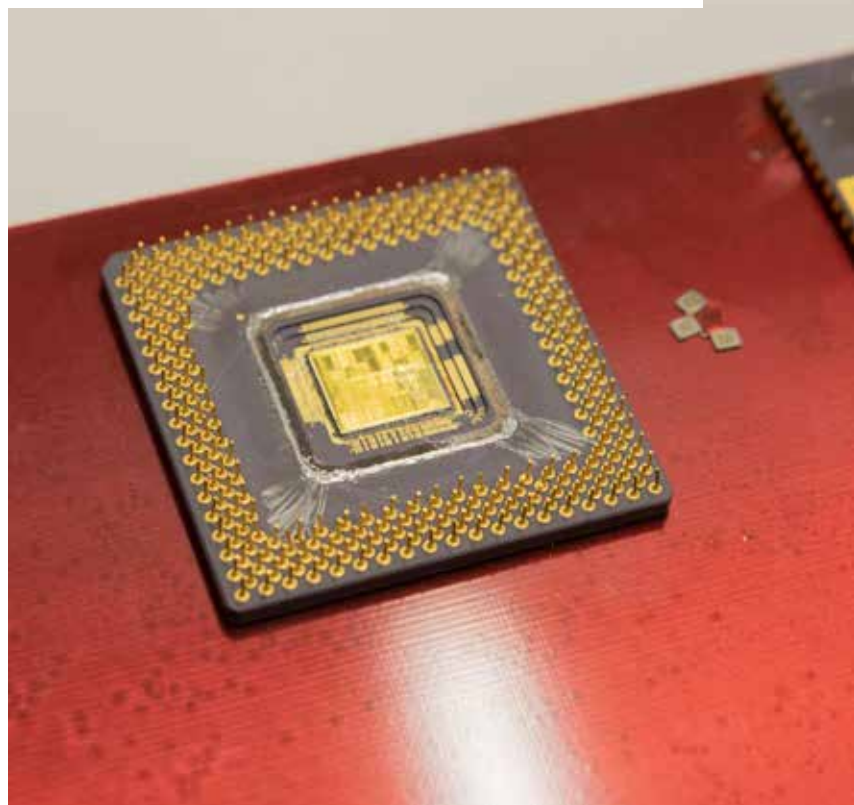
## Protecting convolutional neural networks architectures against cache attacks

Protecting the architecture of neural networks models is paramount to both protect intellectual property and avoid making attacks easier. Convolutional neural networks (CNNs) are the target of several side-channel reverse-engineering attacks aiming at recovering their parameters and hyper-parameters. These attacks often rely on the knowledge of a certain—large—set of hyper-parameters among which the victim model lies. The goal of the potential attacker is then to reduce that search space or even deduce the correct architecture. One such attack, called "Cache Telepathy", monitors access to a common matrix multiplication algorithm (GeMM, the general matrix multiply) in order to determine the victim model's hyper-parameters.

In article [1], we propose to change the order in which the computations are made and add randomness to the said computations in order to mitigate Cache Telepathy. The security analysis of our method shows a multiplication by at least 218 of the reduced search space obtained by Cache Telepathy on VGG-16. This makes our protection effective against this attack.

▸ [1] **H. Chabanne**, **J.-L. Danger**, **L. Guiga**, and **U. Kühne**, "Telepathic Headache: Mitigating Cache Side-Channel Attacks on Convolutional Neural Networks," in *Applied Cryptography and Network Security* (*ACNS*), Kamakura, Japan, Jun. 2021, pp. 363–392.

## Physical unclonable functions robustness

Physical unclonable functions (PUFs) have become an important security primitive, which can greatly enhance authentication mechanisms in digital devices. A good PUF provides a unique identifier, which is not sensitive to memory hacking. However, one of the main drawbacks of PUFs is their reliability, as the bit error rate of their output can be in the range of 3 to 15%. The reliability can be improved if the PUF output contains an indication of the bit reliability. For example, oscillation-based PUFs such as the ring oscillator PUF or the Loop PUF use the sign of a frequency difference as secret, while the difference's magnitude provides reliability information regarding the probability of bit flips. A first approach is to discard unreliable bits. Another approach is the two-metric helper data (TMH) scheme, that does not lose unreliable bits but may be prone to side-channel analysis (SCA) attacks.

In article [2], we proved by experiments that the TMH scheme is vulnerable to SCA. We provide two different approaches for countermeasures that combine modification of the TMH metrics with randomization of the Loop PUF challenge order, and show trade-offs between security and cost. Furthermore, we carry out a deep analysis of the proposed solutions' security, showing that our low-complexity solution, while requiring less randomness, greatly enhances the security against SCA.

▸ [2] L. Tebelmann, **U. Kühne**, **J.-L. Danger**, and M. Pehl. "Analysis and Protection of the Two-Metric Helper Data Scheme," in *International Workshop on Constructive Side-Channel Analysis and Secure Design* (*COSADE*), 2021.
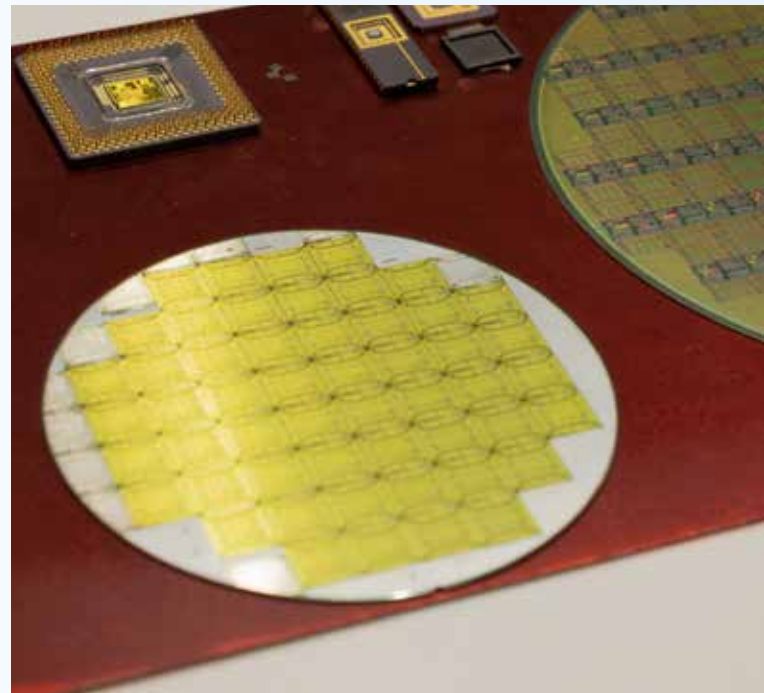
## Microcontroller attack by electromagnetic fault injection

Electromagnetic fault injection attacks represent a threat to cyber-physical systems, being less intrusive and low cost. Developing protection requires a good understanding of the attack mechanisms in order to propose effective countermeasures.

In the thesis [3] defended in July 2021, we address the question of the efficiency of magnetic attacking probes, with a focus on their properties. We present an analytical approach, based on three generic methods, which are used to determine the vulnerabilities of microcontrollers with respect to instructions or data. In addition, whereas usual countermeasures are often based on the assumption that a fault injection implies a single instruction jump, we showed that countermeasures based on a mere instruction-level duplication present vulnerabilities, and proposed some correction mechanisms.

In article [4], we presented earlier some characterization results of intentional electromagnetic interference by measuring its impact within the target, and more precisely on the propagation delay of a combinational logic path. The evaluation of the impact showed that the electromagnetic coupling between the probe and the integrated circuit strongly depended on the properties characterized.

▸ [3] Oualid Trabelsi. "Méthodes pour la modélisation des injections de fautes électromagnétiques." ("Modeling Methods for Electromagnetic Fault Injection") PhD Thesis, *Institut Polytechnique de Paris*, 2021.

▸ [4] **O. Trabelsi**, **L. Sauvage**, and **J.-L. Danger**, "Impact of Intentional Electromagnetic Interference on Pure Combinational Logic," in 2019 *International Symposium on Electromagnetic Compatibility* - EMC EUROPE, Barcelona, Spain, Sep. 2019, pp. 398–403.

▸ See also: **V. Khuat**, **O. Trabelsi**, **L. Sauvage**, and **J.-L. Danger**, "Multiple and Reproducible Fault Models on Micro-Controller using Electromagnetic Fault Injection," Virtuel, France, Jul. 2021.

# ACES

Autonomous Critical Embedded Systems
Systèmes embarqués critiques autonomes

**Group leader**



Laurent Pautet

**Keywords**
Cyber-physical systems, Security and safety, Embedded critical systems, Concurrent and distributed systems, Autonomic systems, Game theory, Blockchain

**Mots-clés**
Systèmes Cyber-Physiques, Sécurité et sûreté, Systèmes embarqués critiques, Systèmes concurrents et répartis, Systèmes autonomiques, Théorie des jeux, Blockchain

**Web**
telecom-paris.fr/aces
**Publications**
telecom-paris.fr/hal-aces

*Non-functional properties of cyber-physical systems such as criticality, autonomy or security must be guaranteed*

The research topics of ACES address **cyber-physical systems** (**CPS**), for which non-functional properties such as criticality, autonomy or security must be guaranteed. These may be systems running in parallel on the same unicore processor, enforcing timing properties as it is the case in real-time embedded systems. They may also consist of distributed systems running on different processors and competitively accessing shared resources in a consistent manner. Similarly, this concurrency can occur in secure systems for which the attacker and the defender are in competition with each other.

To cover the entire continuum of cyber-physical systems, ACES structures its activities around **loosely coupled systems**—distributed services, complex autonomic systems, security in the Internet of things, fault-tolerant and asynchronous distributed computing, model-based testing—and **strongly coupled systems**—real-time systems, deterministic platforms, critical system design process, security and safety, energy consumption of computation. However, this separation remains soft, and the activities overlap and complement each other greatly. For example, critical architectures in the transport domain (mostly categorized as strongly coupled systems) tend to extend to external services—less safe, but more adaptive (mostly categorized as loosely coupled systems).



Permanent members
Membres permanents
**15**

44
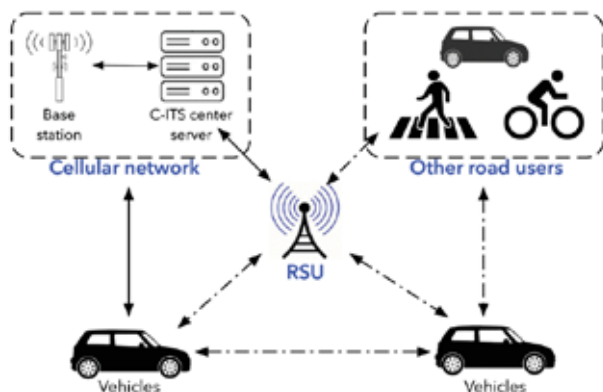
## Secure and safe systems

With the development of modern intelligent transportation system (ITS), reliable, secure and efficient transportation information sharing becomes more and more challenging. In [1], we give a short background tutorial on the main security issues and the different attacks that hinder intelligent transport systems. To enable secure and safe ITS applications, we provide a comprehensive analysis of existing solutions and highlight their strengths and limitations. We also present key challenges in the field, and discuss recent trends that must be factored in by researchers, implementers and car manufacturers, to improve the security of ITS. In the same context, we consider the flash crowd attack in [2]. We propose a novel system that can accurately predict the traffic flow and density in the urban area that can be used to avoid the V2X communication flash crowd situation. By combining the existing grid-based and graph-based traffic flow prediction methods, we use a topological graph convolutional network (ToGCN) followed by a sequence-to-sequence (Seq2Seq) framework to predict future traffic flow and density with temporal correlations. In RFC 8902 [3], we tackle the security of communications in a vehicular environment by extending TLS 1.3 to support a new certificate type, 1609Dot2. We propose several possible scenarios of client and server configurations to ensure more flexibility and interoperability.

▸ [1] A. Lamssaggad, N. Benamar, A. S. Hafid, and **M. Msahli**, "A Survey on the Current Security Landscape of Intelligent Transportation Systems," *IEEE Access*, vol. 9, pp. 9180–9208, 2021.

▸ [2] **H. Qiu**, **Q. Zheng**, **M. Msahli**, **G. Memmi**, M. Qiu, and J. Lu, "Topological Graph Convolutional Network-Based Urban Traffic Flow and Density Prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4560–4569, 2021.

▸ [3] **M. Msahli**, N. Cam-Winget, W. Whyte, **A. Serhrouchni**, and **H. Labiod**, "RFC 8902 TLS Authentication Using Intelligent Transport System (ITS) Certificates," 2020.

## Critical real-time systems

Real-Time considerations and Run-Time mechanisms provide opportunities to tackle issues in various domains.
We address in article [4] the challenge of instantaneous cycle detection (ICD), in the fields of co-simulation and component-based modeling used in cyber-physical systems, where designers import models as building blocks to create a composite model that provides more complex functionalities. The problem becomes indeed harder when the model's input to output dependencies changes for certain events generated internally or externally as inputs. Our method, based on an oracle as a central source of information, is an adequate alternative to the offline analysis methods and significantly reduces the analysis time.
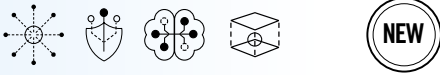In safety-critical systems, many software components of different criticalities or assurance levels need to interact in a timely manner to keep the system and environment safe. Nowadays, these systems are challenged by technological progress resulting in rapid increases in both software complexity and processing demands. In article [5], we consider the mixed-criticality (MC) execution model and homogeneous multi-core processors. We define first a task model incorporating mixed-criticality, real-time and precedence constraints in the form of directed acyclic graphs, followed by a meta-heuristic to solve the scheduling problem of this task model. This meta-heuristic is then generalized to support an arbitrary number of criticality levels. In article [6], we present two contributions on the implementation of global static schedulers for MC systems, aiming to solve the problems of instantaneous migrations and simultaneous mode changes on the multi-core or processor of such global schedulers.

▸ [4] H. Park, A. Easwaran, and **E. Borde**, "Online Cycle Detection for Models with Mode-Dependent Input and Output Dependencies," *CoRR*, vol. abs/2102.01928, 2021.

▸ [5] R. Medina, **E. Borde**, and **L. Pautet**, "Generalized Mixed-Criticality Static Scheduling for Periodic Directed Acyclic Graphs on Multi-Core Processors," *IEEE Transactions on Computers*, vol. 70, no. 3, pp. 457–470, Mar. 2021.

▸ [6] **L. Pautet**, **T. Robert**, and **S. Tardieu**, "Litmus-RT plugins for global static scheduling of mixed criticality systems," *Journal of Systems Architecture*, vol. 118, p. 102221, Sep. 2021.

## ✉ Patent

A method for pairing a first object and a second object that are connected and share a common context. The method comprises receiving, by the first object, over a primary communication channel, a context code being encoded from a context protocol, the context code being transmitted by the second object. The method also includes transmitting, by the first object, over an out-of-band channel, a short authentication string received by the second object. Example embodiments also include a microprocessor-readable storage medium and an apparatus implementing the pairing method.

▸ S. Khalfaoui, A. Villard, and M. Jingxuan, "*Procédé d'appairage*," EP3913951A1

Thesis
Thèse soutenue **1**

Patents
Brevets **3**

Publications
Publications **14**

**45**

# C²

**NEW**

## Group leader

Duong Hieu Phan

### Keywords
Cryptography, Privacy, Network security, Internet of Things, V2X communications security, Coding, Combinatorics, Graphs, Algebraic geometry

### Mots-clés
Cryptographie, Vie privée, Sécurité des réseaux, Internet des Objets, Sécurité des communications V2X, Codage, Combinatoire, Graphes, Géométrie algébrique

### See also
▸ Internet Of Secure Elements, p. 41

### Web
telecom-paris.fr/c2

### Publications
telecom-paris.fr/hal-c2

**Permanent members**
**Membres permanents** — **15**

**Thesis**
**Thèses soutenues** — **3**

**Publications**
**Publications** — **26**

## *New tools and techniques to tackle cryptographic questions arising from new technologies, and fundamental mathematics at the interface between discrete mathematics and computer science*

**C² is a new team, born in October 2021 from the merger of CCN and MC².**

With the deployment of new digital technologies (such as networks, cloud computing, Internet of things, data storage and distribution), the need to master the complexity and security of large digital infrastructures is constantly growing. Communication networks, digital infrastructures, algorithms and data offer an increasingly large surface for exposure to cyber attacks. The risks, in terms of cybersecurity, are multiplying, with an impact on our societies that goes as far as questioning our sovereignty.

The new C² team conducts its research activities in the following domains, including their interactions: **cybersecurity** (mainly inherited from the former CCN team; see last year's review for a full description), **cryptography** and **fundamental research** (mainly inherited from the MC² team).

Regarding the cybersecurity theme, research projects mainly strengthen the previous activities of CCN with new domains such as privacy models, blockchains and protocols for time-sensitive networks. They will be handled in use cases as various as trusted and secure communications in scalable and constrained environments, security data analytics in large scale networks/systems, and attack analysis and countermeasure in information networks.

Cryptography is a fundamental cornerstone of cybersecurity, traditionally supporting data confidentiality, integrity, and authenticity. However, when cryptographic protocols are deployed in emerging applications such as cloud services or big data, the demand for security grows beyond these requirements. Nowadays, data is being extensively stored in the cloud, and users also need to trust the cloud servers/ authorities that run powerful applications. Collecting user data, combined with powerful tools, can come with huge risk in terms of mass surveillance or undesirable data-driven strategies for profit making, while ignoring users' needs.

Our research projects aim to introduce new tools and techniques to tackle cryptographic questions arising from new technologies such as cloud services or big data, and against strong adversaries. The following two directions will be the particular focus of our research: decentralized cryptography to enhance privacy; algorithmic tools for advanced post-quantum cryptographic primitives.

For cybersecurity and cryptography, machine learning is both a valuable tool and an object to study, and quantum technologies form an important framework.
Our last fundamental research theme lies in fundamental mathematics and in areas at the frontier between discrete mathematics and computer science. Based on previous results in algebraic geometry and graph theory, and interacting with new combined research directions, C² explores the following structuring themes: new interfaces in coding theory; new problems in complexity theory and in graph combinatorics.

## Privacy in decentralized cryptography

In a pandemic context, contact tracing (CT) applications are now widely used. However, they remain prone to privacy concerns and financial cost from the user point of view.

Most CT systems are decentralized and impose a significant mobile data cost on end-users because they require them to download a large (several Mbytes), new set of users every day. As the efficacy of CT is proportional to the number of users, minimizing the cost to these users is crucial to the success of CT.

Private set intersection cardinality (PSI-CA) is a secure multi-party computation technique that allows two parties, each holding a set of items, to learn the size of the intersection of those sets without revealing any additional information.

In article [1], we design a modular approach for delegated PSI-CA that is secure against semi-honest parties, where the parties outsource their datasets to an untrusted cloud and let the cloud perform the PSI-CA computation on their behalf. At the end of the computation, the complexity of which is linear in the size of the smaller set $O(n)$, and independent of the larger set's size, the parties only learn the intersection size, while the cloud learns nothing. The main building block of our novel DPSI-CA protocol, which is of independent interest, is the oblivious distributed key pseudorandom function (Odk-PRF). We also design Catalic, a lightweight contact tracing system, on these bases. Catalic provides strong privacy guarantees that can prevent critical attacks (e.g., linkage attacks and false-positive claims). Our experiments show that Catalic is highly scalable, and that the computational and communication costs of our protocol are negligible.

▸ [1] T. Duong, **D. H. Phan**, and N. Trieu, "Catalic: Delegated PSI Cardinality with Applications to Contact Tracing," in *Advances in Cryptology – ASIACRYPT 2020*. ASIACRYPT 2020. *Lecture Notes in Computer Science*, vol 12493. Springer, 2020, pp. 870–899.

## Complexity and graph theory

The decision problems of the existence of a Hamiltonian cycle or of a Hamiltonian path in a given graph, and of the existence of a truth assignment satisfying a given Boolean formula C, are well-known $\mathcal{NP}$-complete problems. Here we study the problems of the uniqueness of a Hamiltonian cycle or path in an undirected, directed or oriented graph, and show that they have the same complexity, up to polynomials, as the U-SAT problem of the uniqueness of an assignment satisfying C. As a consequence, these Hamiltonian problems are NP-hard and belong to the $\mathcal{DP}$ class, like U-SAT.

▸ [2] **O. Hudry**, A. Lobstein, "On the Complexity of Determining Whether There is a Unique Hamiltonian Cycle or Path," to be published in *Journal of Combinatorial Mathematics and Combinatorial Computing*.

▸ See also: **O. Hudry**, A. Lobstein, "Some Rainbow Problems in Graphs Have Complexity Equivalent to Satisfiability Problems," *International Transactions in Operational Research* 29 (3), 2022, 1547-1572.

## Detection of malicious behavior in networks

In Wireless Mesh Networks (WMNs), packet dropping may be due either to an attack or normal loss events such as bad channel quality. Furthermore, in the route discovery phase, path stability is not always considered. We examine in article [3] a special case of denial of service attack in WMNs known as the gray hole attack. In this attack, a node selectively drops some packets which it has to forward along the path. To mitigate this attack, we propose a dropping detection mechanism allowing a mobile node to select the most reliable route to the destination. Our detection module detects misbehaving nodes by comparing the observed packet loss distribution of nodes to the expected ones when they are well-behaved. Among other benefits, our solution aims at decreasing the number of false positives, has a low computational resource requirement on the source node as compared to previously proposed schemes and does not only detect selfish nodes but also forces them to act honestly.

▸ [3] **R. Khatoun**, **Y. Begriche**, L. Khoukhi. "A statistical detection mechanism for node misbehaviours in wireless mesh networks," *IJAHUC* 31(1): 23-35, 2019.

Yes/No

$t_1$

$t_3$

$t_2$

$t = t_1 + t_2 + t_3$
$(m = 3)$
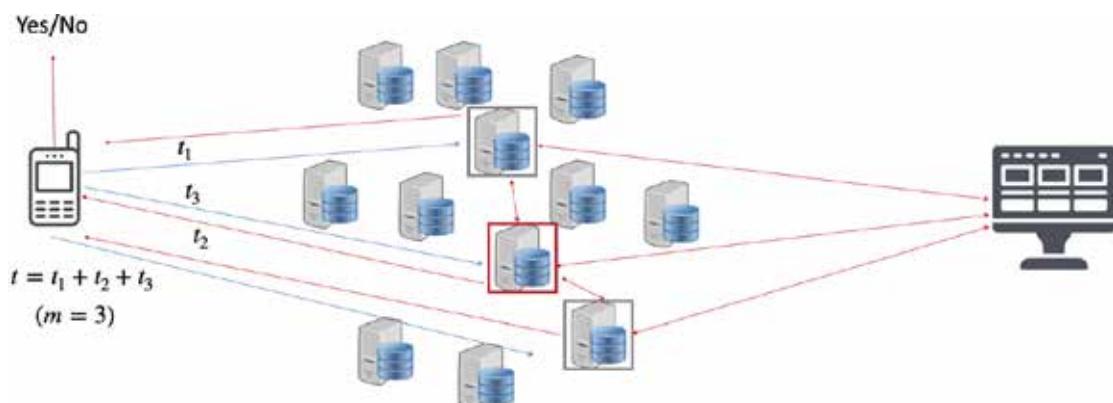
## Neural network pruning and compression

Neural networks need to be efficiently compressed before being deployed to devices with limited storage through tight bandwidth links. Multiple and complementary approaches are possible, to cope with neural networks memory requirements, inference time and energy consumption: 1) redesigning the network topology; 2) pruning: deep architectures are still over-parameterized to be trained effectively, and redundant parameters are pruned at inference time; 3) quantization: representing the parameters and activation functions as fixed-point digits. Most quantization schemes just aim at learning a compressible representation of the parameters rather than properly minimizing the compressed parameters entropy, i.e., obtaining a differentiable formulation to be easily minimized in standard gradient descent-based frameworks.

In article [1] we introduce HEMP, a method that relies on high-order entropy minimization to allow for efficient compression of the parameters of a neural network. The main contribution of HEMP is the differentiable formulation of the quantized parameters' entropy, which can be extended beyond first-order with finite computational and memory complexity and is agnostic to the quantization scheme. We design a regularization term around our entropy formulation that can be plugged into gradient descent frameworks to train a network to minimize the entropy of the quantized parameters. Our experiments proved our approach to be flexible towards non-uniform quantization, scalable towards any entropy order to be minimized and efficient in terms of compression. We also coupled HEMP to LOBSTER, a state-of-the-art pruning strategy which introduces a prior on the weight distribution, which gives a further boost to the final model's compression.

See also previous research on this topic by Enzo Tartaglione, while at University of Turin. Article [2] explores the benefits of structured parameter pruning as part of the MPEG standardization efforts for neural network compression. Article [3] presents a method for learning sparse topologies, exploiting "neural sensitivity"—the contribution of the activity of one given neuron to the network output—as a regularizer.

▸ [1] **E. Tartaglione**, **S. Lathuilière**, **A. Fiandrotti**, **M. Cagnazzo**, and M. Grangetto, "HEMP: High-Order Entropy Minimization for Neural Network Compression," *Neurocomputing*, vol. 461, pp. 244–253, Oct. 2021.

▸ [2] A. Bragagnolo, **E. Tartaglione**, **A. Fiandrotti**, and M. Grangetto, "On the Role of Structured Pruning for Neural Network Compression," in 2021 *IEEE International Conference on Image Processing* (ICIP), 3527–31, 2021.

▸ [3] **E. Tartaglione**, A. Bragagnolo, F. Odierna, **A. Fiandrotti**, and M. Grangetto, "SeReNe: Sensitivity-Based Regularization of Neurons for Structured Sparsity in Neural Networks," *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 1–14.

**Keeping in mind that the sustainable development of both hardware and software does matter for a transition towards a carbon-neutral economy**

## Lightweight Neural Models

Among the problems encountered in Knowledge Base building features entity linking: given, for instance, the mention of an entity on a Web page, map it to any of the predefined entities from a catalog. In article [4], we tackle this problem in the biomedical domain, where linking means mapping biomedical mentions, such as diseases and drugs, to standard entities in a given knowledge base. A specific challenge is that the same biomedical entity can have a wide range of names, including synonyms, morphological variations, and names with different word orders. We show that this problem can be solved by a rather lightweight neural architecture, which needs just a fraction of the parameters of a BERT model and far fewer computing resources. The central idea is to use an alignment layer with an attention mechanism, which can capture the similarity and difference of corresponding parts between candidate and mention names.

▸ [4] **L. Chen**, G. Varoquaux, and **F. Suchanek**, "A Lightweight Neural Model for Biomedical Entity Linking," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 14, pp. 12657–12665, May 2021.

# Pathway to a **sustainable and green IT**

## Optimization through power-aware feature selection

Increasing the battery life of wireless sensors is a main challenge. For certain applications, signal reconstruction is not always required. In this context, Analog-to-feature (A2F) conversion is an acquisition method thought for IoT devices. The aim of this method is to decrease the amount of acquired samples by only extracting useful information directly on the analog signal. Extracted information can be used in a classification task, i.e., to detect voice activity or electrocardiogram anomalies. One of the challenges of A2F conversion is to identify what is useful information for a given application and how to extract it.

In article [5], we studied an A2F converter for ECG anomaly detection based on the non-uniform wavelet sampling. The number of features possibly extracted being very important, we showed how to select a small set of features that provides good classification performance while having a low power consumption.

▸ See also: Energy efficient information acquisition method for IoT, page 29.

▸ [5] **A. Back**, **P. Chollet**, **O. Fercoq**, and P. Desgreys, "Analog-to-feature converter optimization through power-aware feature selection," Bordeaux, France, Oct. 2021.

**See also**
▸ ApproxiNet, p. 91

49

# DIG

## *From theory to software*

**Group leader**

Thomas Bonald

**Keywords**
Data mining, Knowledge bases, Databases, Logic, Graphs, Data streams, Natural language processing, Artificial intelligence, Learning, Cognitive models

**Mots-clés**
Fouille de données, Bases de connaissances, Bases de données, Logique, Graphes, Flux de données, Traitement du langage naturel, Intelligence artificielle, Apprentissage, Modèles cognitifs

**Web**
telecom-paris.fr/dig
**Publications**
telecom-paris.fr/hal-dig

**See also**

Best student paper award at the Extended Semantic Web Conference (ESWC) 2021

**T. Pellissier Tanon** and **F. Suchanek**, "Neural Knowledge Base Repairs," in *European Semantic Web Conference*, Hersonissos (virtual), Greece, Jun. 2021, pp. 287–303.

DIG aims at making data and knowledge easy to extract, process, use and understand. Data here refers to texts, signals (e.g., from IoT devices), graphs (e.g., Web, social networks), log files (e.g., transactions). Knowledge refers to facts and rules.

The research activity of DIG covers both theoretical and practical aspects. On the theoretical side, DIG works on the complexity of query languages for databases and on the analysis of algorithms for large graphs, data streams and time series. DIG also works on cognitive models, like the simplicity theory, seeking to explain the attractiveness of situations or events to human minds.

On the practical side, DIG develops new algorithms for automatically extracting knowledge from text data, new structures and languages for storing and querying data, new machine learning algorithms and new techniques for answering complex questions (e.g., formulated in natural language), for recommending contents, or for detecting trends or anomalies in social media for instance. The team maintains YAGO, a large general knowledge base, and software such as scikit-multiflow, scikit-network and Inkscape.

## Graph mining

DIG has strong expertise in graph mining. Graphs are a near-universal way to represent data. We study mining graphs for patterns and properties. Our particular focus is on the scalability of such approaches.

As part of its activity, the DIG team is engaged in the development of several open source packages. Most notably, we develop and maintain scikit-network, a Python package inspired by scikit-learn for the analysis of very large graphs. The package provides state-of-the-art algorithms for ranking, clustering, classifying, embedding and visualizing graphs. High performance is achieved through a mix of fast matrix-vector products (using SciPy), compiled code (using Cython) and parallel processing. Source code, documentation and installation instructions are available online.

Texts are typically expressed as a sequence of tokens. However, several natural language processing (NLP) problems benefit from a representation of texts using some graph structure: graph between documents, sentences and words, links between documents, knowledge graphs, abstract meaning representation for advanced reasoning. Knowledge graphs and graph neural networks for NLP are part of the topics studied by the DIG team, and further work will be conducted in particular in the joint laboratory TALia for five years, with onepoint launched in June 2021 (see page 85).

## Best paper award at track B of the ICALP (International Colloquium on Automata, Languages and Programming) 2021 conference

The DIG team investigates fundamental problems in theoretical computer science, in the area of data management: how can computers store and query large datasets efficiently? Specifically, DIG studies the management of dynamic data: how to compute information about a dataset, and recompute this information efficiently whenever the dataset is updated. DIG is involved in the Franco-German EQUUS ANR project on this topic. In 2021, it received the best paper award of Track B of the ICALP 2021 conference, one of the top venues in theoretical computer science, for a collaboration with Charles Paperman (Université de Lille).

The focus of article [1] is on validating dynamic textual data: we store a large text that is frequently updated by replacing characters, and we want to quickly check some properties on the text after each update. The properties are expressed using regular expressions (a common pattern matching language), and the article identifies how fast we can recompute whether the properties hold after each update. For instance, we can check that each opening parenthesis ')(' is followed by a closing parenthesis ')' and that there are no nested parentheses. These results could be applied to detect forbidden patterns efficiently in dynamic data, e.g., in computer memory.

▸ [1] **A. Amarilli**, **L. Jachiet**, and C. Paperman, "Dynamic Membership for Regular Languages," in *International Colloquium on Automata, Languages and Programming* (ICALP), Glasgow, Ireland, Jul. 2021.

## Extensions of Knowledge Bases

KBs usually contain mainly binary links between entities—a knowledge representation known as RDF. In the scope of the NoRDF project, the team aims to enrich KBs by beliefs, claims, events, causes, and entire stories (for the reasons behind this project, see the vision paper [2]). The researchers have so far mainly surveyed the existing literature: how to deal with non-named entities [3], how to deal with vague expressions [4], and how to assess whether transformers can reason on natural language [5].

The project "NoRDF: Modeling and Extracting Complex Information from Natural Language Text" is a chair held by Fabian Suchanek (DIG team) and co-held by Chloé Clavel (S²A team). It is funded by the French National Research Agency (ANR), in collaboration with four industrial partners (EDF, BPCE, Schlumberger and Converteo), established for a period of four years, beginning in late 2020. https://nordf.telecom-paris.fr

▸ [2] **F. Suchanek**, "The Need to Move beyond Triples," in *Text2Story – Third Workshop on Narrative Extraction From Texts co-located with 42nd European Conference on Information Retrieval* (*ECIR 2020*), Lisbon, 2020, pp. 95–104.

▸ [3] **P.-H. Paris**, **F. Suchanek**. "Non-named entities—the silent majority," in: *ESWC* short paper track, 2021.

▸ [4] **P.-H. Paris**, **S. El Aoud**, and **F. Suchanek**, "The Vagueness of Vagueness in Noun Phrases," presented at the *International Conference on Automated Knowledge Base Construction* (AKBC), online, United States, 2021.

▸ [5] **C. Helwe**, **C. Clavel**, and **F. Suchanek**, "Reasoning with Transformer-based Models: Deep Learning, but Shallow Reasoning," presented at the *International Conference on Automated Knowledge Base Construction* (AKBC), online, United States, 2021.
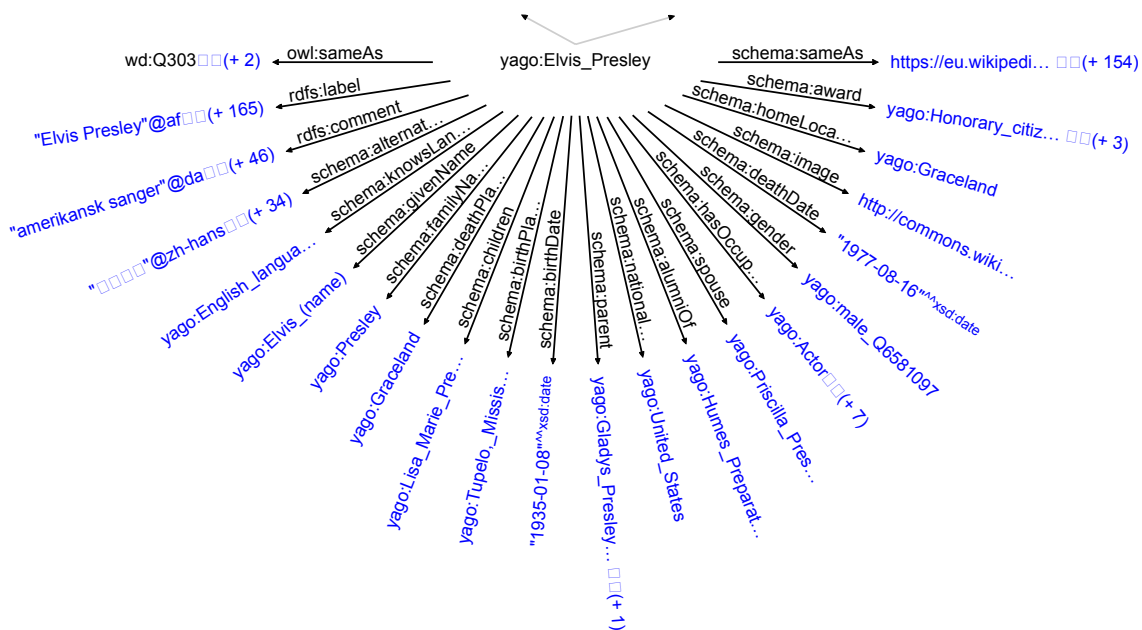
| | | |
|---|---|---|
| Permanent members / Membres permanents | **11** | |
| Thesis / Thèses soutenues | **3** | |
| Patent / Brevet | **1** | |
| Publications / Publications | **12** | |

# DIVA

Design, Interaction, Visualization & Applications
Design, Interaction, Visualisation & Applications

*Novel forms of interaction include wall-sized displays, mixed reality, touch- and gesture-based interfaces*

**Group leader**

Eric Lecolinet

**Keywords**
Human-computer interaction (HCI), Touch- and gesture-based interfaces, Data visualization, Design, Behavior models, Virtual reality, Augmented reality, Mixed reality

**Mots-clés**
Interaction humain-machine (IHM), Interfaces tactiles/ haptiques et gestuelles, Visualisation des données, Design, Modèles comportementaux, Réalité virtuelle, augmentée et mixte

**See also**
▶ ANR projects, p. 54
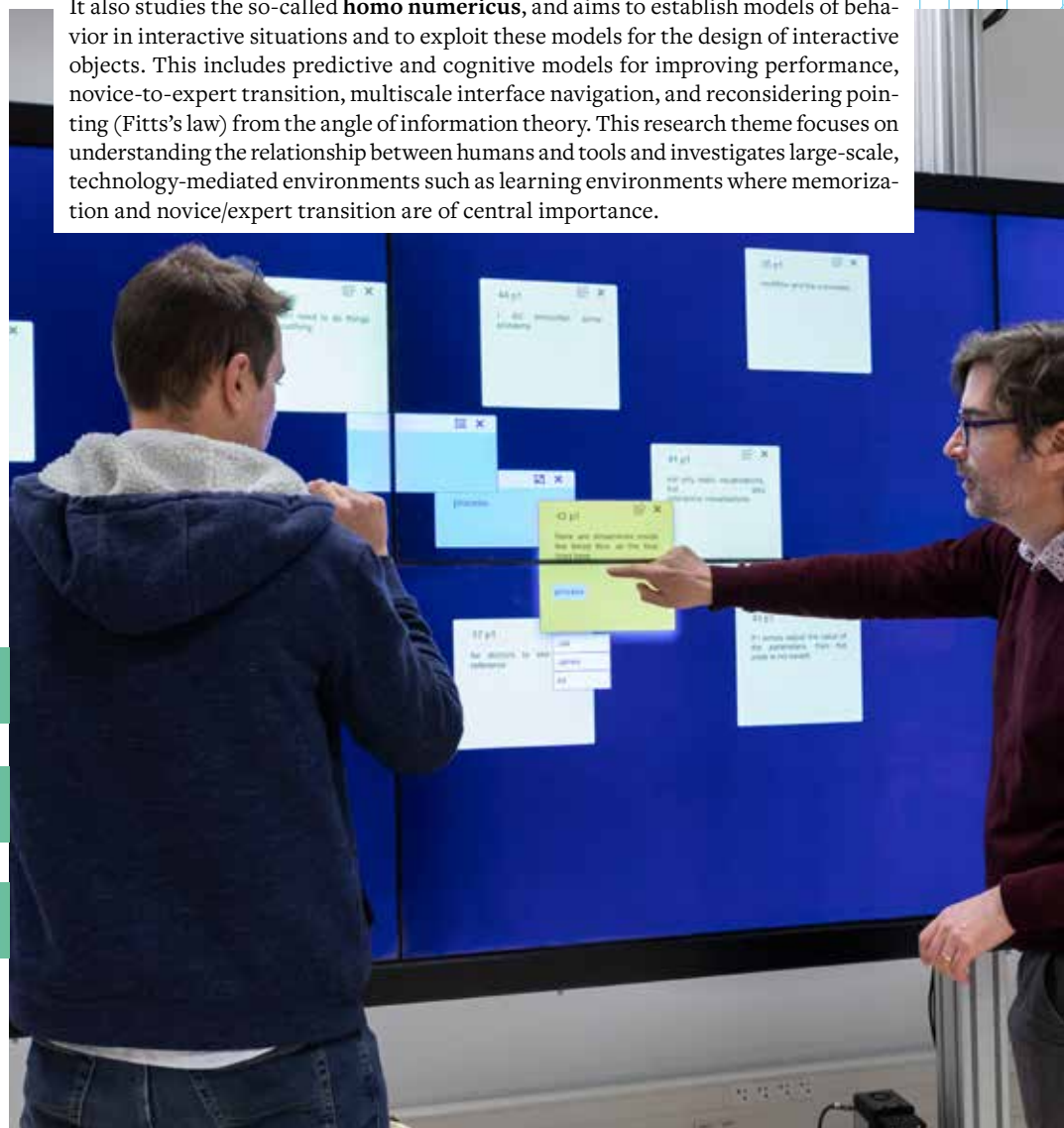
**Web**
telecom-paris.fr/diva
**Publications**
telecom-paris.fr/hal-diva

| | |
|---|---|
| Permanent members Membres permanents | **6** |
| Patents Brevets | **2** |
| Publications Publications | **11** |

DIVA is dedicated to fundamental and applied research on human-computer interaction (HCI), design and data visualization.

DIVA focuses on **novel interaction techniques**, technology mediated environments and on the challenge of constantly representing increasing data volumes, not only on standard computers but also on small, very large, and non-traditional devices. It addresses the complementary aspects of this process by leveraging novel interaction principles and methods, in particular for mixed reality and touch- and gesture-based interfaces, while considering their physical characteristics and studying how they are being used and deployed. It also includes the development of novel interaction and design paradigms, which act as a basis for implementing new interaction techniques for multi-device, multi-surface collaboration and computer science education.

It also studies the so-called **homo numericus**, and aims to establish models of behavior in interactive situations and to exploit these models for the design of interactive objects. This includes predictive and cognitive models for improving performance, novice-to-expert transition, multiscale interface navigation, and reconsidering pointing (Fitts's law) from the angle of information theory. This research theme focuses on understanding the relationship between humans and tools and investigates large-scale, technology-mediated environments such as learning environments where memorization and novice/expert transition are of central importance.

## Potential negative implications of mixed reality technology

Literature in the field of human-computer interaction established the concept of persuasive design which is a practice to influence a user's behavior towards a certain action. While these nudges are mainly framed in a positive direction, recently researchers started to identify and explore instances of persuasive "dark design patterns" which nudge a user towards a behavior which clearly profits the "creator" and not the user. Furthermore, mixed reality devices expand the user's capacity to interact with virtual reality (VR) beyond what would ordinarily physically be possible. For instance, redirected walking and haptics leverage knowledge of the limits of human perception. This, too, by provoking changes in the user's physical movements, leads to perceptibly and imperceptibly nudge their physical actions. We explore the risks posed by the malicious use of interactive devices, and the potential harms that may result: negative impact on the user's physiology (e.g., fitness, eye sight) or immediate physiological harm (hitting an obstacle, running into a wall, hitting a bystander, falling down).

In order to explore this new strand of research, we have proposed and been granted HARMFULVR, an ANR JCJC Young Researcher, (see page 54) project. It aims to create a fundamental understanding of (1) the degree to which a user's physical actions can be influenced in VR using interaction design, (2) the techniques that can be used to nudge a user to perform risky and potentially harmful actions (e.g., collision) and (3) how to design a safety mechanism that is able to protect the user appropriately. In other words, our research is threefold: what is possible? How it is possible? How can we protect the user? Finally, the project also includes a complementary work package that focuses on the ethical implications of immersive persuasive design.

This line of work led to multiple CHI 2022 submissions and acceptances in 2021, with different international teams ("VRception: Rapid Prototyping of Cross-Reality Systems in Virtual Reality", "Consent in the Age of AR: Investigating The Comfort With Displaying Personal Information in Augmented Reality"). In particular, the workshop "Novel Challenges of Safety, Security and Privacy in Extended Reality" is a good reference of work conducted within the HARMFULVR project.
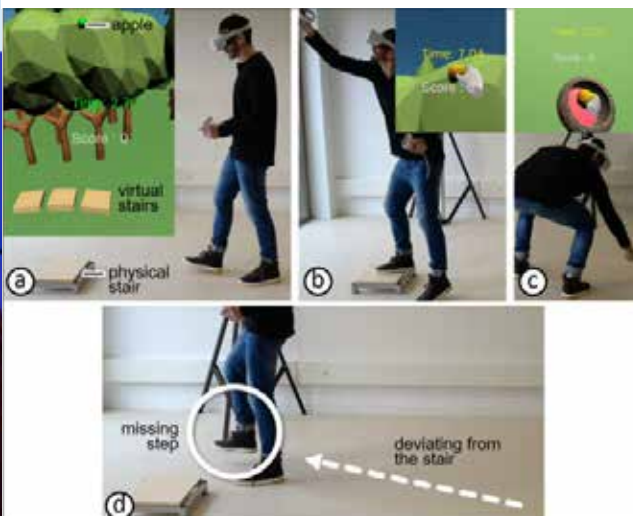


### Thèse

**Jiali Liu**, "Understanding and Supporting Alternatives in Data Analysis Processes", PhD thesis, held on September 23, 2021.

## Shape-changing materials for human-computer interaction

DIVA team has a long track record on research about the touch modality. After a 3-year research program on social touch, further explorations are directing on devices giving realistic touch sensations via a change in their shape (see "SecondSkin", in the "ANR focus" page 54). Tactile feedback can in fact be given via shape-changing interfaces with pneumatic actuation, whereas robotic devices have been used to provide passive (kinesthetic) haptic feedback, especially in VR contexts. However, the constraints of current technology for simulating touch necessitate novel touch devices to be designed, and to test them with different social contexts in mind. Both papers presented below include observational studies and learning about combining and tuning various parameters and various modalities.

As an example of how touch can be used to convey emotions, article [1] proposes using a robotic arm that is augmented with an artificial hand that touches the user's forearm. The primary motivation for choosing this device was its capacity to generate types of touch that are similar to those produced by humans. We excluded vibration motors and focused on technologies providing kinesthetic and tactile feedback that resembles what one feels when touched by the palm and fingers of a human hand. In addition, we also developed Mobilimb, a miniaturized robotic actuator for generating touch, which is shaped as a finger.

In article [2], in a social distancing context, we aim to complement and expand the design space of computer-mediated touch to better support interpersonal communication. Such communications are considered 1) not only for intimate but also semi-intimate relationships between users, and 2) not only for remote but also face-to-face communications. We propose SansTouch, a novel wearable hand sleeve that can be fixed on the top of a smartphone, which embeds inflatable actuators that can produce a realistic touching sensation and also provides visual feedback. This device 1) enables users to exchange bi-directional touch for greetings without physical contacts in face-to-face communications; 2) minimizes the efforts to learn new touch interactions; and 3) can generate a human-like touch sensation.

▸ [1] **M. Teyssier**, G. Bailly, **C. Pelachaud**, and **E. Lecolinet**, "Conveying Emotions Through Device-Initiated Touch," *IEEE Transactions on Affective Computing,* pp. 1–1, 2020.

▸ [2] **Z. Zhang**, **J. Alvina**, R. Héron, **S. Safin**, **F. Détienne**, and **E. Lecolinet**, "Touch without Touching: Overcoming Social Distancing in Semi-Intimate Relationships with SansTouch," in *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 2021.

# Multiple new challenges need to be tackled for obtaining a realistic and persuasive design in virtual/augmented/mixed environments



breaking out
virtual safety boundary

## ANR projects
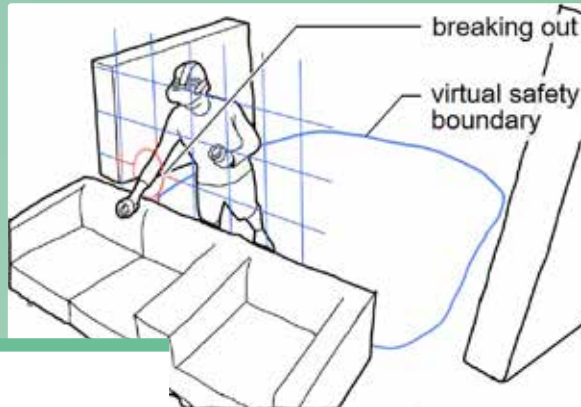
### HARMFULVR

"Assessing the Potential for Physiological Harm of Persuasive Interaction Design in Virtual Reality" (see "Potential negative implications of mixed reality technology", page 53) is one of the 3 ANR (French National Research Agency) projects **DIVA** has been granted in 2021. It is a JCJC (Young researcher) four-year grant, aimed to individual research projects coordinated by a young researcher. It is held by Jan Gugenheimer (see our Young talents, page 10 of the 2019-2020 Annual report).

### SecondSkin

"Shape-Changing" Materials for HCI" (see "Shape-changing materials for human-computer interaction", page 53) is a PRC (collaborative research projects between public entities in a national context) project that follows up on the SocialTouch project presented last year in our previous edition (see Annual report 2019-2020, pp. 46-47). **DIVA** is the project leader, with partners from the **INTERACT** team, physicists from ESPCI Paris - PSL, computer scientists from Laboratoire d'Informatique de Grenoble and an artist and researcher in material science and in interactive arts. The objective is to enrich user experience with the sense of touch through new tactile shape-changing interfaces, both for enabling new interactions and for conveying emotions.

### InterPlay

"Interacting with Rich Display Ecosystems" is another PRC grant which focuses on novel display technologies such as wall-sized displays, very large tabletops, and headsets for virtual reality or augmented reality. The goal of this project is to study rich display environments from a human-computer interaction perspective in order to inform the design of innovative workspaces for crisis management and data analysis. **DIVA** is part of this project, which is led by LISN and also involves IRIT and CEA Toulouse.

# 14 November 2021

### Transitional Interfaces in Mixed and Cross-Reality

In recent years, Head-Mounted Displays (HMDs) became the main form of enabling Extended Reality (XR) applications. However, HMDs come with an inherent problem around exclusion (excluding everyone from the XR experience in the environment) and isolation (isolating the XR user from the environment). In a keynote held during ACM ISS 2021, Jan Gugenheimer presented multiple systems that are redesigning the shape of current HMDs and argued that future XR HMDs have to be designed to be inherently sharable to non-HMD users in the environment.
Hybrid workshop co-organized by Jan Gugenheimer : "Transitional Interfaces in Mixed and Cross-Reality: A new frontier?" ACM ISS 2021. November 14, 2021 in Łódź, Poland

▸ https://idux-jetter.org/ti-iss2021

# Exploring modalities, **novel displays and interaction**

## Computer graphics

In computer graphics, simulating elastic motions of deformable objects is needed in real-time applications such as video games and virtual/augmented reality environments, applications that involve stringent requirements for both stability and efficiency of simulation. The solvers for such applications are often limited to implicit numerical integration methods such as backward Euler, and among such solvers Projective Dynamics (PD) is one of the most popular. However, it suffers from unwanted dissipation of energy and momenta, i.e., artificial damping, which is inherited from the backward Euler scheme. To address this problem, a solution referred to as FEPR has been proposed in recent years. It first takes the full integration steps of the underlying solver and then projects the results onto a constant energy-momentum manifold.

In the below article, we extend the unconstrained minimization problem of PD's time integration into a constrained one with position-based energy-momentum constraints. Our "solver-integrated constraint approach" is distinguished from a "projection-after-solve approach" such as FEPR. The experimental results show that our method produces more visually-appealing elastic motions, faster. Our method effectively and efficiently preserves the important physical quantities (i.e., the total energy and momenta) thus avoiding undesired damping. Through a variety of animation scenarios, we demonstrated that our method yields vivid rotational and deformable motions in real-time. Furthermore, our method enables users to readily create the desired animations by directly controlling the physical quantities to be preserved on the fly.



▸ M. H. Kee, **K. Um**, W. Jeong, and J. Han, "Constrained Projective Dynamics: Real-Time Simulation of Deformable Objects with Energy-Momentum Conservation," *ACM Trans. Graph.*, vol. 40, no. 4, Jul. 2021.

**Group leader**

Romain Alléaume

**Keywords**
Quantum communications, Quantum cryptography, Quantum computing, Quantum key distribution (QKD)

**Mots-clés**
Communications quantiques, Cryptographie quantique, Calcul quantique, Distribution quantique de clé (QKD)

**See also**
▸ Quantum communication infrastructure, p. 7
▸ Quantum secure networks, p. 61

**Web**
telecom-paris.fr/iqa
**Publications**
telecom-paris.fr/hal-iqa

## Theoretical work and demonstrations of quantum communication protocols, based on nonlinear and quantum optics

IQA conducts research on quantum information, quantum Internet, quantum communication and cryptography. Theoretical work is combined with the development of an experimental platform enabling demonstrations of quantum communication protocols to be made, based on nonlinear and quantum optics.

A research theme brings together theoretical and experimental work on **quantum cryptography**, and in particular continuous variable quantum key distribution (CV-QKD). It also consists in addressing challenges aimed at advancing the integration and impact of quantum cryptography within the cybersecurity ecosystem. This work is being carried out as part of the Quantum Technology Flagship and of the EuroQCI initiative led by the European Commission.

A more fundamental aspect of the future quantum Internet is the experimental investigation of new sources of light for quantum communications that are compatible with fiber-optic telecommunications networks.

Another important activity lies at the intersection of computer science and quantum information, with an emphasis on **optical quantum computing**. In view of this, machine-learning-based techniques are being developed to enable differentiable simulations of quantum optical systems. This constitutes an innovative method to optimize parameterized quantum circuits.

IQA also works on frequency-encoded linear optics **quantum computing**, leveraging telecom-optimized optical components. The objective is to build an experimental multimode quantum processing platform to demonstrate applications in quantum computing and quantum cryptography.
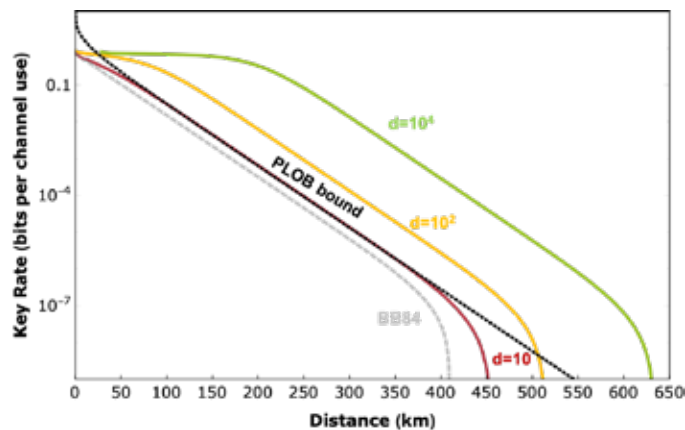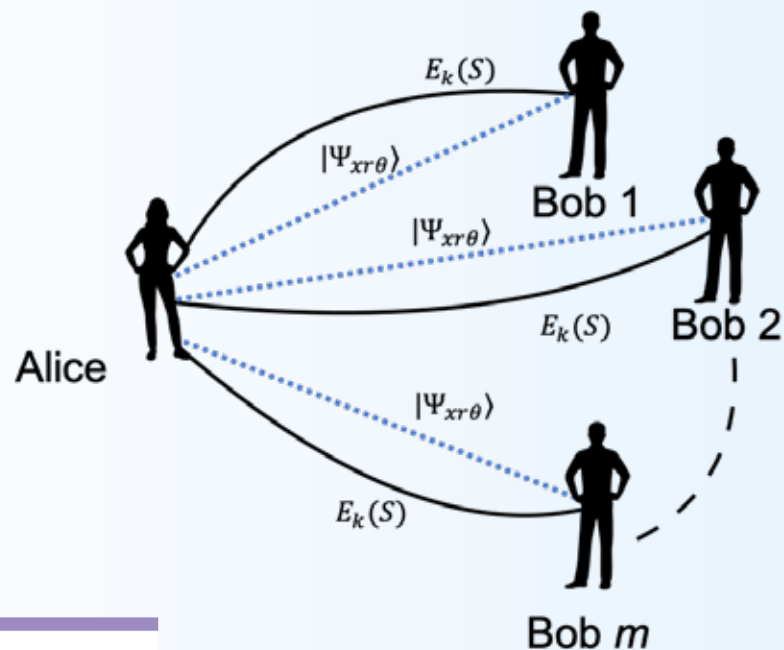
## Vulnerability analysis and security evaluation of quantum cryptographic systems

A necessary condition for the industrial take-off of quantum key distribution (QKD) resides not only in the ability to engineer cost-effective QKD systems, but also in the capacity to provide solid guarantees regarding their security. To guarantee a very high security level for QKD, forward-looking methods and standards in quantum cryptography implementation security will be adopted, following a methodology similar to the one used to certify the security of classical crypto-systems, called Common Criteria (CC). We demonstrate in [1] the practical relevance of this methodology—applicable to any practical quantum cryptographic system, provided it has reached a sufficient maturity level—by conducting vulnerability analysis on an experimental continuous-variable (CV)-QKD system.

We considered two practical methods to mount the saturation attack in CV-QKD. The first strategy relies on inducing detector saturation by performing a large coherent displacement. It corresponds to a high attack rating with regard to the CC, and thus a less pressing threat. A second attack strategy simply consists in saturating the detector with an external laser. It appears to constitute a primary threat for practical CV-QKD systems. Using the CC along with a metric called attack potential we introduced QKD to highlight the benefits of combining theoretical security considerations with vulnerability analysis based on attack ratings, in order to guide the design and engineering of practical QKD systems towards the highest possible security standards.

▸ [1] R. Kumar, **F. Mazzoncini**, H. Qin, and **R. Alléaume**, "Experimental Vulnerability Analysis of QKD Based on Attack Ratings," *Scientific Reports*, vol. 11, no. 1, p. 9564, May 2021, doi: 10.1038/s41598-021-87574-4.
▸ See also, video: https://2021.qcrypt.net/sessions/contributed2a

## Quantum computational timelock

On the journey to QKD technology development, researchers and engineers pursue goals that are structured around two dimensions: practicality (how to build efficient and cost-effective QKD systems) and security (how to guarantee an effective security gain with respect to existing classical techniques). We explore in [2] a complementary space: considering security models weaker than unconditional security and characterizing the gain in practicality (i.e., performance and functionality, over cost). This approach requires a clear bench-marking of *the security gain*, with respect to classical cryptography, and the security cost.

We introduce the Quantum Computational Hybrid (QCH) security model, where we assume that computationally secure encryption may only be broken after a much longer time than the coherence time of available quantum memories. We propose an explicit $d$-dimensional key distribution protocol, called MUB-quantum computational timelock (MUB-QCT), where one bit is encoded on a qudit state chosen among $d+1$ mutually unbiased bases (MUBs).

We demonstrate that MUB-QCT enables everlasting secure key distribution—an authenticated key establishment scheme has the everlasting security property if it is secure against adversaries that have unlimited computational power after the protocol execution—with input states containing up to $O(\sqrt{d})$ photons. Implementable with a large number of modes with current or near-term multimode photonics technologies, this leads to important improvements and has the potential to provide a radical shift in the performance and practicality of QKD: the ability to operate securely between one sender and many receivers, whose implementation might moreover not be trusted; significant performance increase (key rates and attainable distance over fiber).

▸ [2] **R. Alléaume**, and **N. Vyas**, "Everlasting Secure Key Agreement from the Quantum Computational Timelock," *The International Conference on Quantum Communications,* Paris, France, Oct. 2022.

# RMS

**Group leader**

Marceau Coupechoux

**Keywords**
Mobile networks, Cellular networks, 5G, 6G, Future Internet, Internet of things, Smart Grids, Data centers, Network storage, Cloud and edge computing, Virtualization (SDN/NFV), Network management and control, Metrology, SDR, Performance analysis, Optimization, Mathematical models (stochastic geometry, game theory, algebric topology), Algorithm design, Artificial intelligence

**Mots-clés**
Réseaux mobiles, Réseaux cellulaires, 5G, 6G, Internet du futur, Internet des objets, Réseaux électriques intelligents, Centres de données, Informatique en nuage et en périphérie, Stockage de données, Virtualisation (SDN/NFV), Contrôle et gestion des réseaux, Métrologie, SDR, Analyse de performances, Optimisation, Modélisation mathématique (géométrie stochastique, théorie des jeux, topologie algébrique), Conception d'algorithmes, Intelligence Artificielle

**See also**
▸ Monitoring and AI-based performance optimization for high-speed networks, p. 12
▸ 5G massive Machine Type Communications, p. 60

**Web**
telecom-paris.fr/rms
**Publications**
telecom-paris.fr/hal-rms

## Envisioning the world with 5G and beyond

The research interests of RMS lie in the design and study of future and emerging networking technologies, architectures and services. Research efforts are concentrated around three main themes: **cloudify networking and virtualization, mobile networks and communications, and the Future Internet/Internet of Things**.

The research activities carried out in the RMS group cover a wide range of topics in the general area of networking, ranging from lower layers—access, resource allocation, energy harvesting, radio handover, radio mobility management, data transmission reliability and infrastructure resilience—to higher layers—routing, Information-Centric Networking (ICN), Network Functions Virtualization (NFV), network automation and programmability, edge computing, data centers and cloud architectures, network storage, cloud RAN, IoT networking and service discovery, IP mobility, embedded systems, network telemetry and metrology including quality of user experience.

RMS has a strong expertise in protocol and architecture design. Indeed, RMS is involved in standardization activities (IETF, ETSI) as well as in the design of major research platforms (LISP, FIT, SILECS) for the networking community. Theoretical aspects are tackled using advanced approaches and tools, such as stochastic geometry, simplicial homology, game theory, optimal control and scheduling. RMS is engaged in structural cooperation with international laboratories (Europe, China, India, MENA region). Its expertise is regularly sought by national regulatory bodies and public policy makers. Research activities are also undertaken in the context of national (ANR, FUI) and international research projects (European projects), in close cooperation with major industrial groups in the field.

RMS has joint research laboratories including NewNet@Paris (with CISCO), SEIDO Lab (with EDF), LINCS (with Nokia, Inria, Sorbonne Université and IRT SystemX), and ValaDoE (with Region Pays de la Loire, Nantes Métropole, ENEDIS, and Mines Saint-Étienne).
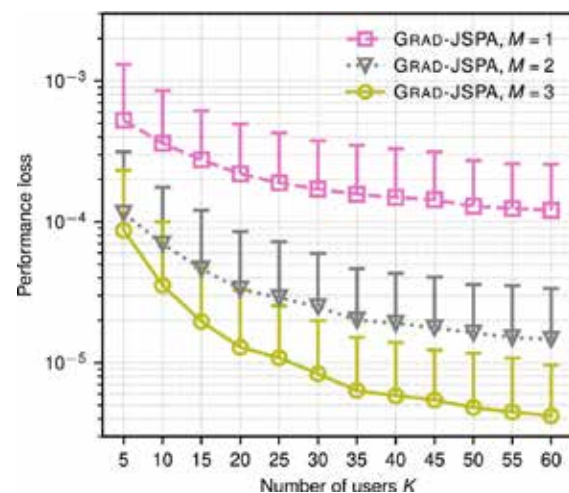
| Permanent members<br>Membres permanents | **15** |
| Thesis<br>Thèses soutenues | **3** |
| Patents<br>Brevets | **15** |
| Publications<br>Publications | **8** |



## Optimization and resource allocation for NOMA

Employing stochastic geometry (SG) to investigate dimensioning for Non-orthogonal multiple access (NOMA) scheme in uplink direction becomes indispensable for 5G and beyond. This comes with challenges that need to investigate uplink served user point process distribution. We prove in [1] that it admits homogeneous Poisson distribution. Secondly, uplink SG models of two schemes are derived to obtain expressions of total required radio block (RB) numbers. We also give their upper bounds to simplify network deployment. Thirdly, simulation results confirm the homogeneous Poisson assumption, and the effectiveness of SG models in determining RB provisions.

Joint subcarrier and power allocation problems (JSPA) in NOMA are NP-hard to solve in general. In article [2], we improve the computational complexity of existing single-carrier power control and user selection schemes. These improved procedures are then used as basic building blocks to design new algorithms. OPT-JSPA computes an optimal solution with lower complexity than current optimal schemes in the literature. $\varepsilon$-JSPA, a fully polynomial-time approximation scheme, stands out by allowing a tight trade-off between performance guarantee and complexity to be controlled. Finally, GRAD-JSPA is a heuristic based on gradient descent. Numerical results show that it achieves near-optimal results with much lower complexity than existing optimal methods.

▸ [1] B. Liu, **P. Martins**, **L. Decreusefond**, **J.-S. Gomez**, and R. Song, "Uplink Dimensioning Over Log-Normal Shadowing for OMA and NOMA Schemes," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 5126–5130, 2021.

▸ [2] **L. Salaün**, **M. Coupechoux**, and C. S. Chen, "Joint Subcarrier and Power Allocation in NOMA: Optimal and Approximate Algorithms," *IEEE Transactions on Signal Processing*, vol. 68, pp. 2215–2230, 2020.

In 2021, Lou Salaün, was awarded the first PhD thesis prize of the IDIA Department (Computer Science, Data and Artificial Intelligence, Institut Polytechnique de Paris), and the PhD thesis award of the Fondation J Hadamard Programme Gaspard Monge, for his work on resources allocation in NOMA.

▸ Lou Salaün, "Resource Allocation and Optimization for the Non-Orthogonal Multiple Access," PhD Thesis, Institut Polytechnique de Paris, 2020.
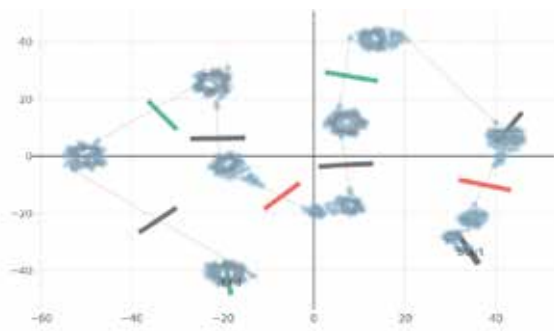
## Distributed services and telemetry for cloud-native networking

Model-driven telemetry enables the real-time collection of hundreds of thousands of counters on large-scale networks, each describing a particular aspect of device behavior. Labeling the data in such a high dimensional dataset would imply the annotation of all the counters by a domain expert, a process not only time consuming if at all possible, but also too subjective and dependent on every expert point of view. In article [3], we propose a method to find an intelligible selection of operational counters, i.e., that can be understood and analyzed by a human. It is a general method for event description in multivariate time-series using both data behavior, through the change score, and contextual information, through information retrieval on the feature names only. In one tested dataset, it generates selections of less than ten highly meaningful features out of more than six thousand, for events with different root causes.

In article [4], a broader set of counters, not limited to the key performance indicators, is explored with an unsupervised approach. The goal is to leverage the dependencies between the counters in order to discover complex state changes that might have otherwise slipped the operator's view. DESTIN is an online multivariate change detection methodology based on principal angles between subspaces, which is well-suited to detect temporal state changes in noisy operational data of a network element, to assess the state changes of an element. Results show that DESTIN performs better than state-of-the-art solutions for almost all types of changes. Due to the robustness in preprocessing assumptions, it performs better for traffic blackholing and interface shutdown changes.
This work has been carried out as part of the Cisco NewNet@ Paris chair, and with the collaboration of École Polytechnique for [4].

▸ [3] T. Feltin, **P. Foroughi**, W. Shao, F. Brockners, and T. H. Clausen, "Semantic feature selection for network telemetry event description," in *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, 2020, pp. 1–6.

▸ [4] **P. Foroughi**, W. Shao, F. Brockners, and **J.-L. Rougier**, "DESTIN: Detecting State Transitions in Network elements," in 2021 *IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 2021, pp. 161–169.

# Path to the Internet
# of the Future

## 5G massive Machine Type Communications

The 5G-mMTC project aims to develop the software and hardware tools necessary for the implementation of a 5G IoT system, based on static and moving communicating objects, and interoperable with the existing 4G RAN. It will allow manufacturers to test innovations, new use cases and associated business models in real configuration and on a large scale. Developed as part of the French national recovery and resilience plan, the project's objective is to consolidate national sovereignty in the key area of connected objects.

Two representative use cases will be implemented during this 36 months project. The first one, with the French Cycling Federation, will enable real-time analysis of athletes' data and their performances. The objective is to build a private 5G network, deployable anywhere in the world.

The second use case, with Électricité de France, will enable real-time management and telemetry of an entire fleet of existing heterogeneous sensors (from sensitive infrastructure to near-customer equipment). Here the slice technology of the 5G core network is used to differentiate the services offered by the sensors.

In both cases, data sensitivity is very high and data leakage must absolutely be avoided. Therefore, the security aspect of the 5G solution will be studied in depth to prevent and eliminate cyber attacks.

Télécom Paris' expertise in 5G will be mobilized via the RMS team and its two experiment rooms: the FIT IoT Lab and the 5G/4G SDR Lab.

# Integrating present and future networks, allowing new usages, exploring novel concepts

## Quantum secure networks

Télécom Paris is part of QSAFE (Quantum Network System Architecture for Europe), an EU-backed consortium (27 partners, 5 of them being mainly industrial) which delivered in December its interim results, aimed at paving the way for the implementation of the European Quantum Communication Infrastructure (EuroQCI).

Within the "Conceptual Security Architecture" task, IQA analyzed and compared security services, based on QKD and post quantum cryptography (PQC) that can be deployed with technology available in the near-term and provide a clear security advantage. The analysis focuses on secure communications schemes based on QKD for which the choice of encryption algorithm is the central design choice that strongly influences the security gain, the performance (communication rate vs. distance), and the practicality and cost (required infrastructure). The two main options are one-time pad (OTP) encryption that offers information-theoretic security for SC (ITS SC), unachievable with algorithmic cryptography, but with the constraint that the secure communication rate will be limited to the QKD rate, and AES encryption (or more generally any symmetric encryption algorithm based on a computationally secure one-way function).

From a security gain viewpoint, the choice of QKD+OTP is the stronger option, well adapted to use cases related to data with long-term-security requirements and for which relatively low data rates are acceptable. This should be considered as a priority in the context of EuroQCI, in order to provide security services with the strongest security gain. It can, for example, be used for high-security key transport and more generally high-security key management, possibly over a large network, with trusted nodes. In such use cases, QKD+OTP can be thought of as a special (information-theoretically and automatized) form of trusted courier.

QKD key establishment, combined with AES Encryption, QKD+AES, can be used to perform secure communication at high-data rates (e.g., with 100 Gbit/s AES encryptors). However, as pointed out by ANSSI [ANSSIQKD], the situation in terms of security gain is less favorable in that case. The IQA team has however analyzed and characterized QKD+AES, offering a clear security advantage and a stronger protection against key leakage, when used over point-to-point links where it can be combined with existing algorithmic key exchange mechanisms to increase the overall security.

▸ **R. Alléaume**, "Conceptual Security Architecture Interim Report, QSAFE: Detailed Study for the European Quantum Communication Infrastructure," Télécom Paris, Contract, Dec. 2021.

▸ [ANSSIQKD] https://www.ssi.gouv.fr/en/publication/should-quantum-key-distribution-be-used-for-secure-communications

## Mid-infrared opto-electronics for high-speed data transmission

Free-space optical communication (FSO) is desirable in applications where fiber deployment is quite difficult and expensive, or when the emitter and receiver need to be mobile, or as a substitution network after disasters such as earthquakes, and for private communications.
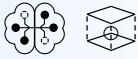
Article [1] investigates the potential for direct modulation of two types of quantum cascade devices, namely quantum cascade lasers (QCLs) and interband cascade lasers (ICLs) for high-speed transmission, both lasers emitting in one of the atmospheric transparency windows near 4 μm. It provides guidelines for future work on the implementation of a real-field experiment with mid-infrared quantum cascade lasers.

Our study shows that off-the-shelf packaged QCLs are compatible with high-speed long-distance transmissions with realistic optical attenuation. With an NRZ format, data rates up to 546 Mbits/s were achieved and with an RZ format, the rate was increased to 680 Mbits/s. The results also showed the first free-space communication with a room-temperature ICL. The data rate of 110 Mbits/s obtained with an NRZ format was strongly improved by selecting an RZ format. This sets the maximum transmission speed at 300 Mbits/s and is promising for future experiments with these energy-efficient optical sources.

In article [2], we show that two unidirectionally coupled QCLs operating in the chaotic regime and the synchronization between them allow for the extraction of the information that has been camouflaged in the chaotic carrier, hence providing a straightforward tool for privacy directly on the physical layer. We achieved the first proof-of-concept of a secure communication operating at a wavelength of 5.7 μm, with a message encryption at a bit rate of 0.5 Mbit/s.

▸ [1] **O. Spitz**, **P. Didier**, **L. Durupt**, D. A. Díaz-Thomas, A. N. Baranov, L. Cerutti, **F. Grillot**, "Free-Space Communication With Directly Modulated Mid-Infrared Quantum Cascade Devices," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 28, no. 1, pp. 1–9, 2022.

▸ [2] **O. Spitz**, A. Herdt, J. Wu, G. Maisons, M. Carras, C.-W. Wong, *et al.*, "Private communication with quantum cascade laser photonic chaos," *Nature Communications*, vol. 12, no. 1, p. 3327, Jun. 2021.

# IMAGES — Image, Modeling, Analysis, GEometry, Synthesis
Image, Modélisation, Analyse, GEométrie, Synthèse

**Group leader**

Yann Gousseau

*In-depth knowledge of special application fields and applications with a societal impact*

**Keywords**
Mathematics and artificial intelligence for images, Stochastic image modeling, Machine learning and deep learning, Image and video restoration, Image and video synthesis and editing, Discrete mathematics, Algebraic and structural models, Symbolic AI, Computer graphics, Geometrical and topological analysis, Rendering, Simulation, Remote sensing imaging, SAR imaging, Medical imaging, Computational photography, Creative industries

**Mots-clés**
Mathématique et intelligence artificielle pour l'image, Modélisation stochastique des images, Apprentissage, Restauration d'images et de vidéos, Synthèse et édition d'images et de vidéos, Mathématiques discrètes, Modèles algébriques et structurels, IA symbolique, Informatique graphique, Analyse géométrique et topologique, Rendu, Images et données de télédétection, Imagerie SAR, Imagerie médicale, Photographie computationnelle, Industries créatives

**Web**
telecom-paris.fr/images

**Publications**
telecom-paris.fr/hal-images

IMAGES carries out research in image analysis and understanding, computer vision and 3D computer graphics, covering the entire spectrum of the modeling, analysis, transformation, representation, interpretation and synthesis of images, 3D and digital objects. It focuses on the **modeling of images**, tri-dimensional and numerical objects, with the development of **mathematical models**, ranging from the physical acquisition to the high-level interpretation, and **artificial intelligence models** (logic, knowledge representation, spatial reasoning, machine learning). It also has strong expertise in **computer graphics** for geometric modeling, image synthesis, animation and simulation, virtual reality and 3D interactive systems.

IMAGES's strength is the mastery of the whole chain of **image and data processing**, ranging from acquisition steps to **information extraction** and **higher level understanding**, along with the distinctive characteristic of covering many aspects of AI: recent machine learning methods, including deep learning, as well as symbolic artificial intelligence. It also benefits from the strong interactions between all research topics—for instance, computer graphics, image processing and image analysis.

IMAGES' expertise relies on in-depth knowledge of special application fields—medical imaging, remote sensing imaging, computational photography, interactive games and creative industries—and applications with a societal impact, thanks to strong and long-term collaborations with academic and state or industrial partners.
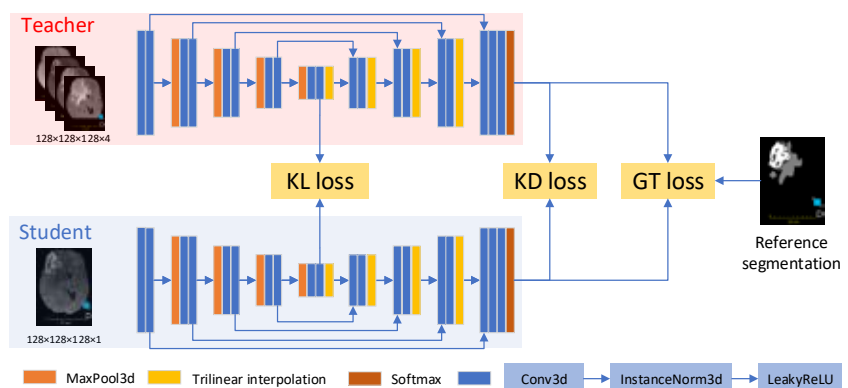
## Remote sensing and SAR imaging

Speckle is a very strong noise affecting synthetic aperture radar (SAR) images acquired for earth observation. Deep learning approaches are very efficient to reduce it but it is necessary to develop semi-supervised or auto-supervised strategies because no "ground truth" images are available. In recent works we have proposed using multi-temporal series, now widely available, to train efficient despeckling networks or to use the complex nature of the signal to derive an auto-supervised approach with single images.

This last work offers promising perspectives since huge archives of complex SAR data could be used in an unsupervised way to train efficient speckle reduction networks. Further research includes extension of the MERLIN unsupervised framework to multi-temporal data, and interferometric and polarimetric modalities.

▸ [1] **E. Dalsasso**, L. Denis, and **F. Tupin**. "As If by Magic: Self-Supervised Training of Deep Despeckling Networks with MERLIN," *IEEE Transactions on Geoscience and Remote Sensing* Institute of Electrical and Electronics Engineers, 2022, 60.

▸ [2] **E. Dalsasso**, L. Denis, and **F. Tupin**. "SAR2SAR: A Semi-Supervised Despeckling Algorithm for SAR Images," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 2021.
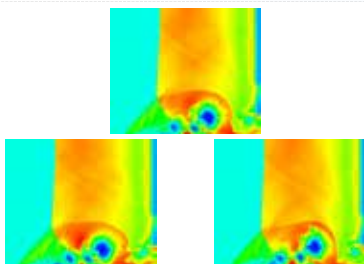
| | |
|---|---|
| Permanent members Membres permanents | **14** |
| Thesis Thèses soutenues | **2** |
| Patents Brevets | **6** |
| Publications Publications | **44** |

## Medical imaging

Traditional supervised learning with deep neural networks requires a tremendous amount of labeled data to converge to a good solution. For 3D medical images, it is often impractical to build a large homogeneous annotated dataset for a specific pathology. Self-supervised methods offer a new way to learn a representation of the images in an unsupervised manner with a neural network. In particular, contrastive learning—finding similar and dissimilar things in a self-supervised manner—has shown great promises by (almost) matching the performance of fully-supervised CNN on vision tasks. Nonetheless, this method does not take advantage of available meta-data, such as participants' age, viewed as prior knowledge.

Article [3] proposes to leverage continuous proxy metadata, in the contrastive learning framework, by introducing a new loss called $y$-Aware InfoNCE loss. Specifically, we improve the positive sampling during pre-training by adding more positive examples with similar proxy meta-data with the anchor, assuming they share similar discriminative semantic features. With our method, a 3D CNN model pre-trained on 104 multi-site healthy brain MRI scans can extract relevant features for three classification tasks: schizophrenia, bipolar diagnosis and Alzheimer's detection. When fine-tuned, it also outperforms 3D CNN trained from scratch on these tasks, as well as state-of-the-art self-supervised methods. We also made a step towards a debiased algorithm by demonstrating that our model is less sensitive to the site effect than other SOTA fully supervised algorithms trained from scratch.

Our code is available at https://github.com/Duplums/yAwareContrastiveLearning

Acquiring multiple modalities is usually not possible in a clinical setting due to a limited number of physicians and scanners, and to limit costs and scan time. In article [4], we propose KD-Net, a novel framework to transfer knowledge from a trained multi-modal network to a mono-modal one. We illustrate its effectiveness in brain tumor segmentation. However, the proposed framework stays generic and can be applied to any encoder-decoder segmentation network. The gain in segmentation accuracy and robustness to errors makes it highly valuable for real-world clinical scenarios where only one modality is available at test time.

We proposed in article [5] a machine learning subtype discovery method that aims at finding relevant homogeneous subgroups with significant statistical differences in a given class or cohort. Subtype discovery consists in finding interpretable and consistent subparts of a dataset, which are also relevant to a certain supervised task. Our method is generic, it can incorporate any clustering method and can be driven by both binary classification and regression. It has been tested on synthetic toy examples, MNIST and a neuro-psychiatric dataset.

Codes and examples are available in a scikit-learn-compatible Python package. https://github.com/neurospin-projects/2021_rlouiset_ucsl

▸ [3] **B. Dufumier**, **P. Gori**, J. Victor, A. Grigis, M. Wessa, P. Brambilla, *et al.*, "Contrastive Learning with Continuous Proxy Meta-Data for 3D MRI Classification," Strasbourg (virtuel), France, Sep. 2021.

▸ [4] **M. Hu**, **M. Maillard**, Y. Zhang, **T. Ciceri**, **G. La Barbera**, **I. Bloch**, *et al.*, "Knowledge Distillation from Multi-Modal to Mono-Modal Segmentation Networks," in *MICCAI*, Lima, Peru, Oct. 2020, vol. LNCS 12261, pp. 772–781.

▸ [5] **R. Louiset**, **P. Gori**, **B. Dufumier**, J. Houenou, A. Grigis, and E. Duchesnay, "UCSL : A Machine Learning Expectation-Maximization framework for Unsupervised Clustering driven by Supervised Learning," 2021.

# MM

## Group leader

Stéphane Lathuilière

*Deep networks, large-scale training and generative models are being investigated as promising methods for video compression and generation*

### Keywords
Image and video compression, Transport and orchestration of multimedia content, Immersive media and interactivity, Deep learning for image and video generation, Deep learning with limited supervision

### Mots-clés
Compression images et vidéo, Transport et orchestration des contenus multimédia, Médias immersifs et interactivité, Apprentissage profond pour la génération d'images et de vidéos, Apprentissage avec supervision limitée

### See also
▶ Stéphane Lathuilière, p. 19
▶ Neural network pruning and compression, p. 48

### Web
telecom-paris.fr/mm

### Publications
telecom-paris.fr/hal-mm

MM's research concerns the entire life cycle of multimedia documents and signals—acquisition, coding, transmission, transport, interactivity and recognition. It can be divided into three main areas: compression and transmission of images and videos, interactive service delivery and orchestration, deep learning for image and video compression and processing. The latter is an emerging topic of research and will be incorporated into the first one in the long term.

Traditionally active in **image and video compression**, MM focuses recent projects on emerging **immersive formats** such as multiview and 360 video, high-dynamic-range video and digital holography. Research activities are carried out in the field of robust video transmission, with a consistent set of contributions in the domains of distributed video coding, network coding and, more recently, linear video coding. In addition to these activities, MM continues to work on more traditional but still very relevant problems such as rate-distortion optimization, transcoding, and super-resolution.

Immersive visual formats require new solutions not only for compression, but also for transmission architectures, which is why MM also works on interactive and adaptive **multimedia streaming**, and on multimedia **orchestration**—using multiple sources and multiple networks to deliver the stream to multiple screens. Furthermore, **deep networks** and large-scale training are being investigated as a promising method for video compression. Generative models could also be used to this end, in addition to their capacity to generate images and videos.

Permanent members
Membres permanents

**7**
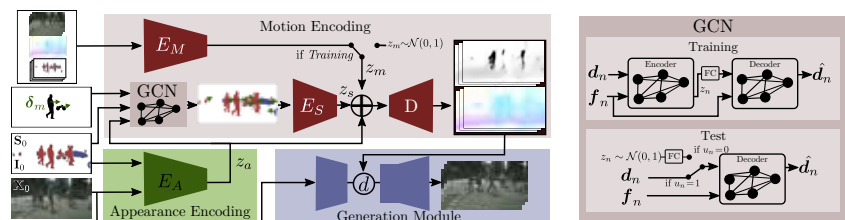
## Lightly-supervised learning and domain adaptation

Novel class discovery (NCD) consists in training a network to simultaneously classify a set of labeled classes while discovering new ones in an unlabeled dataset. Advantage is taken of the supervision available on the labeled set that allows enriched representations of the images to be learnt, transferring this know-how to discover unknown classes in the unlabeled set. During training, the data is divided into a set of labeled images and a set of unlabeled images, with the classes assumed to be disjoint. These two sets are used to form a single network to classify the known and unknown classes.

Motivated by the need to simplify NCD approaches, and inspired by recent advances in self-supervised learning, we propose in article [1] to unify all objectives in a single loss function. Our method differs from the literature in that we use pseudo-labels in combination with ground truth labels according to a unified objective (UNO) that allows for better cooperation and less interference between supervised and unsupervised learning. We also removed the need for costly self-supervised pretraining, making NCD more practical. The effectiveness of the proposed approach is demonstrated through extensive experiments and careful analysis. UNO outperforms the state-of-the-art on several benchmark sets (CIFAR-100, ImageNet).

▸ [1] E. Fini, E. Sangineto, **S. Lathuilière**, Z. Zhong, M. Nabi, and E. Ricci, "A Unified Objective for Novel Class Discovery," 2021.

▸ See also [2], where actions are learned from unlabeled data under self-supervised learning.

(a) Playable Video Generation          (b) Our results



1) The user clicks the destination
4) Video is generated
2) Objects' interactions are modelled
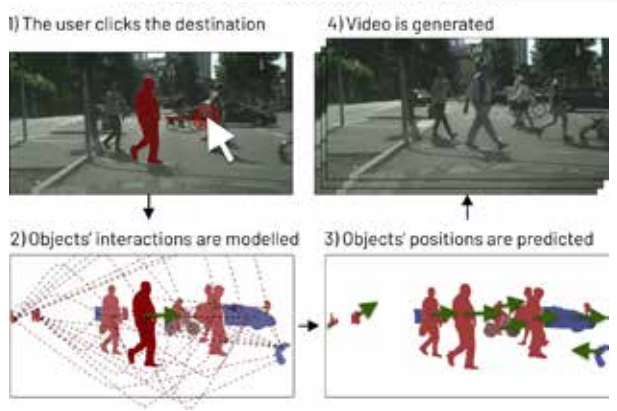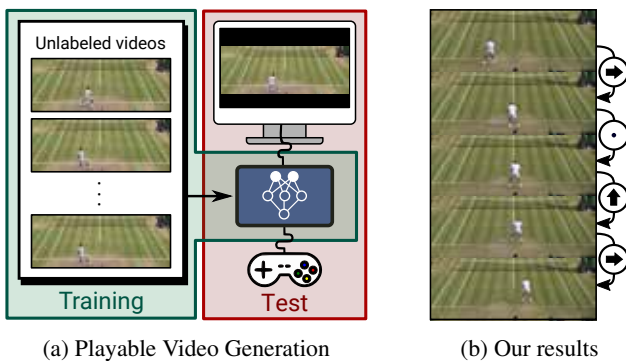3) Objects' positions are predicted

## Image and video generation using deep neural networks

From an early age, humans are able to identify key objects, and how each object can interact with its environment. This skill, developed in an unsupervised way, is particularly noticeable when watching sports or video games, where one is able to understand and anticipate actions without ever being given an explicit list of plausible actions, to analyze the technique, as well as the "amount" of action left to perform.
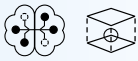
We study in article [2] the unsupervised learning problem of playable video generation (PVG). Our goal is to allow a user to control a generated video by selecting a discrete action at each time step, as in a video game. The goal is to learn a set of discrete actions from real-world video clips in an unsupervised manner in order to offer the user the ability to generate new videos interactively. The difficulty of the task lies both in learning semantically consistent actions and generating realistic videos conditioned on the user input. We propose a novel framework for PVG, named Clustering for Action Decomposition and DiscoverY (CADDY). We use an encoder-decoder architecture where a discrete bottleneck layer is employed to obtain a discrete representation of the transitions between consecutive frames. The network is forced to learn a rich action space using, as main driving loss, a reconstruction loss on the generated video. We demonstrate the effectiveness of the proposed approach on several datasets in a wide variety of environments.

Code and examples are available at https://willi-menapace.github.io/playable-video-generation-website

In a similar context of user-controlled video generation, we propose in article [3] "Click to Move", a new approach to generate complex video scenes. Here, the user selects objects in a scene and specifies their final location by mouse clicks. Our system uses the initial frame and its associated segmentation map to compute the motion representations for generating the image sequence. The underlying deep network incorporates a graph convolution network modeling the interaction between objects, to infer the motion of all objects in the scene.

Code is available at https://github.com/PierfrancescoArdino/C2M

▸ [2] W. Menapace, **S. Lathuilière**, S. Tulyakov, A. Siarohin, and E. Ricci, "Playable Video Generation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (*CVPR*), Jun. 2021, pp. 10061–10070.

▸ [3] P. Ardino, M. De Nadai, B. Lepri, E. Ricci, and **S. Lathuilière**, "Click to Move: Controlling Video Generation with Sparse Motion," *International Conference on Computer Vision*, Online, France, Oct. 2021.

| Thesis Thèses soutenues | **3** | Patents Brevets | **26** | Publications Publications | **13** |

# S²A

**Group leader**

Stephan Clémençon

**Keywords**
Statistics, Probabilistic modeling,
Machine learning, Data science, Audio
and social signal processing

**Mots-clés**
Statistique, Modélisation probabiliste,
Apprentissage machine, Science des
données, Traitement des signaux audio
et sociaux

**See also**
▸ Non-smooth stochastic
optimization, p. 35

**Web**
telecom-paris.fr/s2a
**Publications**
telecom-paris.fr/hal-s2a

| | |
|---|---|
| Permanent members<br>Membres permanents | **23** |
| Thesis<br>Thèses soutenues | **17** |
| Publications<br>Publications | **73** |

S²A structures its research around four interconnected topics: **Probabilistic modeling, simulation and mathematical statistics, Machine Learning and optimization, Audio data analysis and signal processing, Social computing**. Work is conducted on a wide variety of data—high-dimensional, structured, heterogeneous / multiscale, massive, incomplete / censored data, data streams, audio signals, text, weak signals and extreme and rare events—and use cases—including anomaly detection, brain data analysis and emotion recognition via handwriting or EEGs, audio data analysis, signal processing, recommender systems and graph inference, chatbots, opinion-mining and human-agent interaction, ranging from social signal perception to generation. This data, which comes from the Internet of Things or from distributed file systems, is often massive, needing massively parallelized / distributed computation, large scale learning and real time processing.

In addition to these constraints and parameters, a key to efficient data processing explored by S²A, and a major problem in mathematics, is to find a sparse, adaptive representation of information and to devise algorithms to calculate it quickly. From an operational point of view, embedding predictive models should not compromise the autonomy of the systems they rely on, and should take the constraints of (nearly) real time operations into account. This raises questions on compression information and the rules for processing it. Moving up to the socio-technical aspects, S²A's work also considers the level of delegation to be granted to smart systems, as well as questions of reliability, interpretable decisions, privacy preservation and ethics. These high level considerations are taken into account by S²A to develop statistical learning techniques that hold up even if part of the data has been contaminated, due to biases in measurements or the deliberate intention to impair the operation of an automated system.

## ① Best paper award at ACM ICMI 2021, Montréal

The ACM ICMI 2021 conference, which was held in a hybrid form in Montréal, Canada, from 18-22 October 2021, was the premier international multidisciplinary research forum on multimodal human-human and human-machine interaction, interfaces and system development. The conference addressed the theoretical and empirical foundations, technologies and techniques of multimodal processing that define the field of multimodal interaction analysis, interface design and system development.

The paper "Exploiting the Interplay between Social and Task Dimensions of Cohesion to Predict its Dynamics Leveraging Social Sciences" has been granted best paper award. It was carried out as part of the ANR project GRACE (ANR-18-CE33-0003-01) led by Giovanna Varni and funded under the French Artificial Intelligence national plan.

▸ [1] **L. Maman**, **L. Likforman-Sulem**, M. Chetouani, and **G. Varni**, "Exploiting the Interplay between Social and Task Dimensions of Cohesion to Predict Its Dynamics Leveraging Social Sciences," in *Proceedings of the 2021 International Conference on Multimodal Interaction*, New York, NY, USA: Association for Computing Machinery, 2021, pp. 16–24.

**See also**
▸ Social Computing, p. 69

## Multiview speech enhancement

Speech enhancement aims to recover clean speech from a noisy signal. In case of a single-channel input, the noise reduction is often accompanied by an increase in speech distortion. Multichannel speech enhancement can overcome this limitation by exploiting the spatial information provided by several microphones. In [2], we extend a previously introduced distributed DNN-based time-frequency mask estimation scheme that can efficiently leverage spatial information captured in so-called compressed signals which are pre-filtered target estimations.

We obtain competitive performance under realistic acoustic conditions. We show that the nodes in the microphone array cooperate by taking advantage of their spatial coverage in the room and we give insights into the potential benefit of sending noise estimation rather than target estimation.

▸ [2] N. Furnon, R. Serizel, **S. Essid**, and I. Illina, "DNN-Based Mask Estimation for Distributed Speech Enhancement in Spatially Unconstrained Microphone Arrays," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 29, pp. 2310–2323, 2021.

Video to Music retrieval

Music to Video retrieval
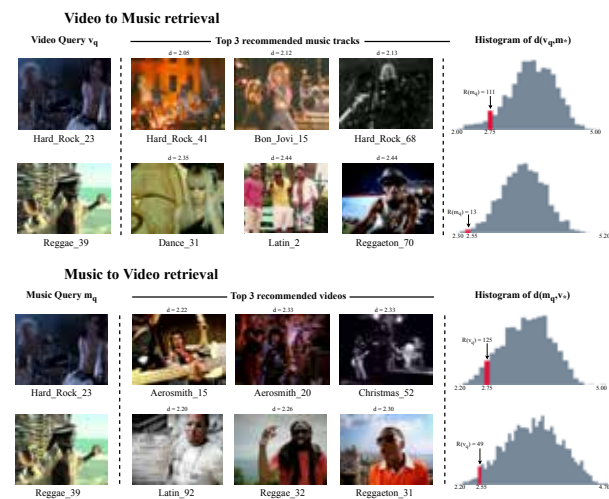
Qualitative results on the MVD for Video to Music and Music to Video recommendation tasks. All distances were multiplied by 1,000 for readability.

*From high-dimensional, structured, heterogeneous | multiscale, massive, incomplete | censored data and signal processing to social computing and human-robot interaction*
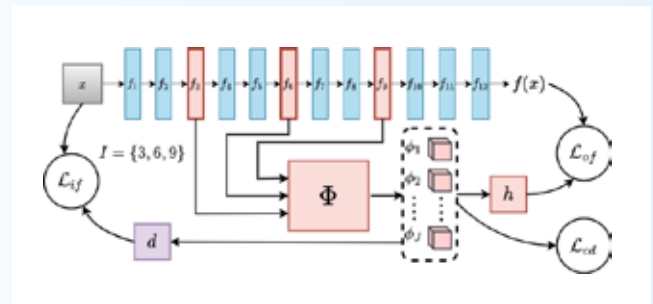


## Video-to-music recommendation

Recommending a video automatically given a music track, or vice versa, has become an important asset for the audiovisual industry. However, while both music and video have specific temporal organizations, most current works do not consider these specificities and only focus on making a global recommendation of a medium.

As a first step toward the improvement of these recommendation systems, we study in [3] the relationship between music and video temporal organization, with the assumption that the movement in the music is correlated to that in the video. We validate this hypothesis with several internationally recognized music video experts. We then perform a large-scale analysis of official music-video clips using music information retrieval description and computer vision tools. Our study confirms that a "language of music-video clips" exists; i.e. editors favor the co-occurrence of music and video events using strategies such as anticipation. It also highlights that the amount of co-occurrence depends on the music and video genres.

Another step is based on multimodal learning paradigms, specifically for linking audio data to corresponding video data. Our objective is to answer questions such as: "Given a query video clip, which music track from a database is most suitable to serve as a soundtrack? Likewise, given a music track, which video best illustrates its content?". In [4], we build upon a recent video-music retrieval system (the VM-NET), which originally relies on an audio representation obtained by a set of statistics computed over handcrafted features. We demonstrate that using audio representation learning from several open-source audio embeddings largely improves recommendations. We also validate the use of the cross-modal triplet loss originally proposed in the VM-NET compared to the binary cross-entropy loss commonly used in self-supervised learning.

▸ Best Video Award ISMIR 2021 [3] **L. Prétet**, **G. Richard**, and **G. Peeters**, "Is There a "Language of Music-Video Clips" ? A Qualitative and Quantitative Study," presented at the *International Society for Music Information Retrieval Conference*, Virtual Event, France, Nov. 2021.

▸ [4] L. Prétet, G. Richard, and G. Peeters, "Cross-Modal Music-Video Recommendation: A Study of Design Choices," in *International Joint Conference on Neural Networks, IJCNN 2021, Shenzhen, China, July 18-22, 2021*, 2021, pp. 1–9.

## Explainable AI

In an attempt to bring semantic understanding to a very different domain, we look into explaining the decisions of a black box machine learning model with the help of several decision trees. Indeed, deep learning models need to be human-interpretable, if we want them to be used in any situation where the user or citizen needs some level of understanding of the algorithmic prediction. One popular family of approaches builds surrogate models—simple and interpretable models that mimic the behavior of the black box model. In article below, we propose to mimic a given black box classifier, not by a single surrogate model, but by several—one for each class. Our user study confirms that the metrics we proposed, confidence and generality, are important features of interpretation, and that users prefer our interpretations over others.

▸ **N. Radulovic**, **A. Bifet**, and **F. Suchanek**, "Confident Interpretations of Black Box Classifiers," in *2021 International Joint Conference on Neural Networks (IJCNN)*, 2021, pp. 1–8.

# How to build trustworthy relationships with our artifacts?

# Human-centered AI

## Social computing

Studying social computing and human-robot interaction provides opportunities to address various challenges in multilingual and business contexts, using specific or synthetic conversational datasets. A crucial step is the identification of underlying information in the user's utterance: communicative intent or dialogue acts, and emotions.

In [1], we introduce new pretraining losses tailored to learn generic multilingual spoken dialogue representations. Indeed, spoken dialogue systems need to be able to handle both multiple languages and multilingualism inside a conversation. The goal of our losses is to expose the model to code-switched language. In order to scale up training, we introduce a new method to automatically obtain several million conversations with multilingual input context in different languages. This was derived from OpenSubtitles, a huge multilingual corpus composed of 24.3G tokens. We test the generic representations on MIAM, a new benchmark composed of five dialogue act corpora on French, Italian, English, German and Spanish, as well as on two novel multilingual tasks (i.e., multilingual mask utterance retrieval and multilingual inconsistency identification). We demonstrate that our loss leads to better performances on downstream tasks, that involve both monolingual conversations and multilingual input conversations.

In the business world, many companies might benefit from studies dedicated to more precise environments such as after sales services or customer satisfaction surveys. In [2], we focus on data from a live chat support in which we aim to detect emotions and their evolution in the conversational flow. Multiple challenges arise that range from exploiting restricted, small and mostly unlabeled datasets to finding and adapting methods for such a context. We tackle these challenges by using few-shot learning and assuming that it can serve conversational emotion classification for different languages and sparse labels. We propose a variation of prototypical networks for sequence labeling in conversation, tested on two datasets with different languages: daily conversations in English and customer service chat conversations in French. When applied to emotion classification in conversations, our method proves to be competitive even when compared to other ones.

A new line of research around the study of the user's trust dynamics during human-robot interaction began in late 2020. Users expect a robot to act in a realistic manner to create a more human-like relationship. In this context, trust is an essential concept as it determines the effectiveness of the system and its acceptance by users. The novelty of the proposed idea relies on building computational models of the dynamics between trust-related multimodal social signals of the user and the multimodal social signals of the robot using sequence mining algorithms.

The understanding of trust dynamics in human-robot interaction (HRI) is still low and systematic studies of multimodal trust-related behaviors in HRI are relatively rare given the rising popularity of the topic. To bridge this gap, we present in [3] a novel coding system TURIN (Trust in hUman Robot INteraction) to study trust in HRI. A preliminary assessment of the coding system was carried out on the Vernissage dataset. Results show a significant agreement between expert annotators.

▸ [1] **P. Colombo**, **E. Chapuis**, **M. Labeau**, and **C. Clavel**, "Code-Switched Inspired Losses for Spoken Dialog Representations," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic, Nov. 2021, pp. 8320–8337.

▸ [2] **G. Guibon**, **M. Labeau**, H. Flamein, L. Lefeuvre, and **C. Clavel**, "Few-Shot Emotion Recognition in Conversation with Sequential Prototypical Networks," in *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, Online and Punta Cana, Dominican Republic, Nov. 2021, pp. 6858–6870.

▸ [3] **M. Hulcelle**, **G. Varni**, **N. Rollet**, and **C. Clavel**, "TURIN: A Coding System for Trust in hUman Robot INteraction," in *2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, 2021, pp. 1–8.

# i3

## interdisciplinary institute on innovation

i3 has more than 200 members, including over 70 professor-researchers and more than 100 PhD students. The high-level research produced by the institute combines academic excellence with relevant findings for research users.

The LTCI and i3 labs frequently work together on topics that call on their respective disciplines. Projects include research on shape-changing materials for human-computer interaction (p. 54) and on the explainability of artificial intelligence to combat money laundering and the funding of terrorism, which forms the object of a research chair (XAI-4-AML, p. 11), as part of the Operational AI Ethics project. In 2017, the High Council for Evaluation of Research and Higher Education (HCERES) assessed i3 and awarded it the highest grade for the quality of its research.

You will find in the following pages the presentation of the three Télécom Paris teams participating in i3, alongside researchers from Mines ParisTech and École Polytechnique.

i3 covers a large range of disciplines, from economics, management, sociology, psychology and ergonomics to information and communication science. As such, it offers a number of multidisciplinary seminars and research days throughout the year.

The INTERACT team organizes a monthly interdisciplinary seminar (Sociology, Psychology, Ergonomics, Information and Communication Sciences) focusing on the uses of information and communication technologies. The approaches put forward are both concerned with observing and understanding these uses "in action" and bringing together different theoretical approaches on activity and communication.

"*Les sciences sociales sur un plateau*" is a regular seminar, initiated by the SID team and organized around the (digital and) social sciences, alternating presentations by social science researchers from Télécom Paris and other institutions in the Saclay area. Sandrine Bubendorff and Caroline Rizza (SID) held a seminar on the use of Wikipedia during the Covid-19 pandemic in October 2020 (see page 77).

The Economics of Digitization Seminar is organized by Digital Economics Paris, a joint initiative of Télécom Paris, IMT Business School and Université Paris Saclay. It is co-led by Ulrich Laitenberger (ECOGE).

The Summer School - IP Paris / i3 Doctoral Seminar "Renewing the view on innovation: a multidisciplinary approach" took place in Aussois from July 5 to 9, 2021. Multiple perspectives on researchers' work were presented on the themes of "crises, risks and uncertainties", and also "innovation perspective", "creation and creativity" and "data and artificial intelligence".

A day was organized on September 22, 2021, to bring together the various studies carried out in i3 on the theme of mobilities, and to establish a dialogue beyond disciplinary frameworks on a topic that involve rethinking the relationships between societies, innovation and the environment. Béatrice Cahour and Marie Hoarau presented their work on the shared robot-taxis (see page 75).

The Interdisciplinary Institute on Innovation (i3) was created in 2012 by Mines ParisTech and Télécom Paris. École Polytechnique joined it in 2014. On January 1, 2015, i3 became a CNRS Mixed Research Unit (UMR 9217), which comprises five research teams working on innovation and social sciences. The Télécom Paris Economics and Social Sciences Department (ECOGE, INTERACT, SID) constitutes one of those five teams, together with the École Polytechnique Management Research Center (CRG) and the CERNA, CGS and CSI teams at Mines ParisTech.

Innovation is a multidimensional phenomenon and central to the way societies operate. A proper understanding of its multiple dimensions requires exchanges among a range of disciplines, including economics, management, sociology, law, design and the psychology of ergonomics. The institute's researchers deal with fundamental issues that bring together very active international research communities and are also highly relevant to companies and decision makers.

Research projects are organized into four cross-disciplinary themes, covering the entire field of innovation. Each theme offers an effective forum for exchanges among disciplines and among the institute's various teams. Transformation in innovative companies; Design theories and models; Innovation regulation; Innovation practices, take-up and democratization. With its research and training activities, i3 helps meet today's major challenges: the digital world, healthcare, innovation, energy and sustainable development.

Research is usually carried out in partnerships involving companies, not-for-profits or government agencies. The institute also heads the work of several research and training chairs. It engages in discussions with companies, not-for-profit organizations and public authorities using innovative exchange platforms.
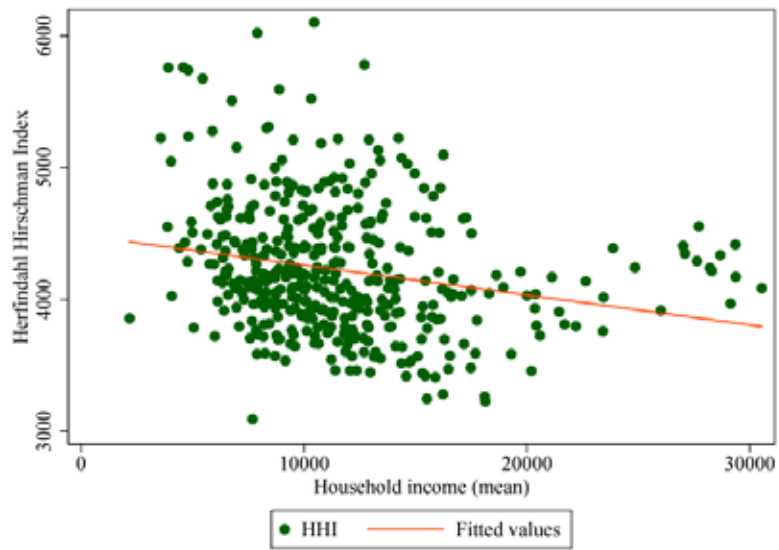
# ECOGE

## Group leaders

Marc Bourreau

David Massé

The ECOGE team is evolving in 2022. It now brings together the economists of Télécom Paris, and becomes the Economics team (ECO). This new ECO team is headed by Marc Bourreau.

## Keywords
Digital platforms, Data, Innovation, Telecommunications, Infrastructure, Strategy, Digital transformation, Public policy, Regulation, Competition policy, Intellectual property, Privacy, Industrial organization, Econometrics, Qualitative methods

## Mots-clés
Plateformes numériques, Données, Innovation, Télécommunications, Infrastructures, Stratégie d'entreprise, Transformation numérique, Politique publique, Régulation, Politique de la concurrence, Propriété intellectuelle, Vie privée, Économie industrielle, Économétrie, Méthodes qualitatives

## See also
▶ Ulrich Laitenberger, p. 18

## Web links to the new team
telecom-paris.fr/ecoge
## Publications
telecom-paris.fr/hal-ecoge



ECOGE is a multidisciplinary group that focuses on economics, law and management and whose research projects explore our understanding of digital transformations and their impact on the economy, companies and public policy. Research activities are based on a wide range of quantitative and qualitative methodologies and, in keeping with the group's mission, research is both theoretical and empirical, and intended for academics as well as for public bodies. For instance, ECOGE has designed a theoretical data taxation model, which has contributed to France Stratégie's work on that topic. Research projects provide fresh insight on topics that include digital platforms, blockchain, artificial intelligence, algorithms, data, the sharing economy, digital entrepreneurship, frugal innovation and data privacy.

Some of the research is conducted in partnership with companies or public institutions, as part of the activities carried out by six research chairs. In this way, theories can be developed, discussed and redefined, with the assurance that the knowledge produced is relevant and of a high scientific standard.

*Exploring our understanding of digital transformations and their impact on the economy, companies and public policy*

| | |
|---|---|
| Permanent members Membres permanents | **15** |
| Thesis Thèses soutenues | **1** |
| Publications Publications | **44** |

## Digital platforms, algorithms, and AI

Digital technologies are pro-competitive because they reduce search costs and other costs, but they can also have anti-competitive effects. For example, data brokers may sell data strategically to soften competition between their business clients [1]. Streaming platforms such as Spotify have an incentive to bias their recommendation system to increase their market power *vis-à-vis* content providers [2].

A relevant question is whether the Internet has made markets more competitive overall. We provided evidence that price differences across countries have remained similar online and offline, suggesting that the Internet has not made international markets more integrated [3].

▸ [1] **D. Bounie**, **A. Dubus**, and **P. Waelbroeck**, "Selling Strategic Information in Digital Competitive Markets," 2021, *RAND Journal of Economics*, 52(2), pp. 283-313.

▸ [2] **M. Bourreau**, and G. Gaudin, "Streaming Platform and Strategic Recommendation Bias," 2021, *Journal of Economics & Management Strategy*.

▸ [3] N. Duch-Brown, **L. Grzybowski**, A. Rohman, and F. Verboven, "Are Online Markets More Integrated than Traditional Markets? Evidence from Consumer Electronics," 2021, *Journal of International Economics*, vol. 131, 103476.

## Competition and regulation in digital markets

Competition in digital markets is limited by different market characteristics. Consumers may face high switching costs when moving to another supplier, creating consumer inertia, as we have shown for the smartphone market [4]. On the supply side, minority shareholding in rival firms may soften competition [5].

Various policies aim at enhancing competition in digital markets. Deregulating markets and allowing entry may bring substantial benefits to the consumers, as illustrated by the entry of Free in the French mobile market [6]. However, the benefits of these policies can be unevenly distributed. For example, in South Africa, wealthier households benefited more from the liberalization of the telecommunications sector than poor households [7].

▸ [4] **L. Grzybowski**, and A. Nicolle, "Estimating Consumer Inertia in Repeated Choices of Smartphones," 2021, *Journal of Industrial Economics*, vol. 69(1), p. 33-82.

▸ [5] S. Heim, K. Hüschelrath, **U. Laitenberger**, and Y. Spiegel, "The Anticompetitive Effect of Minority Share Acquisitions: Evidence from the Introduction of National Leniency Programs," 2021, *American Economic Journal: Microeconomics*.

▸ [6] **M. Bourreau**, Y. Sun, and F. Verboven, "Market Entry, Fighting Brands and Tacit Collusion: Evidence from the French Mobile Telecommunications Market," 2021, *American Economic Review*.

▸ [7] R. Hawthorne, and **L. Grzybowski**, "Distribution of the Benefits of Regulation vs. Competition: The Case of Mobile Telephony in South Africa," 2021, *International Journal of Industrial Organization*, vol. 74, 102673.

▸ <u>Student Paper Award</u>: **Raphaela Andres** and Olga Slivko (2021), "Combating Online Hate Speech: The Impact of Legislation on Twitter," *ZEW Discussion Paper* No. 21-103, Mannheim.

▸ <u>CRESSE & CPI Awards for Young Researchers - Best Digital Economy Paper by the CCIA</u>: **Michelangelo Rossi**, "Competition and Reputation in a Congested Marketplace: Theory and Evidence from Airbnb", 2021.

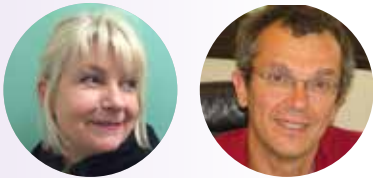## Digital transformations for workers and companies

In the past few years, every single industry, company, and worker has been affected by the digital revolution. Several ECOGE projects have explored this phenomenon, which has produced radical change in many companies' organization and processes and more broadly, in work practices. For instance, research shows that IT companies pay developers who work on free software projects, which has led to new hybrid forms of work [8]. Another study examines how algorithm-based monitoring systems would be analyzed under European fundamental rights law and, in particular, the Court of Justice of the European Union's (CJEU's) case law on the processing of personal data for the purpose of fighting crime [9]. Lastly, Internet platforms often provoke a form of reconfiguration of the different markets in which they operate. An article explores the role and impact of AI on the strategic dynamics of the platforms [10].

▸ [8] M. O'Neil, **L. Muselli**, M. Raissi and **S. Zacchiroli**, "'Open Source Has Won and Lost the War: Legitimising commercial–communal hybridisation in a FOSS project," *New Media & Society*, 2020, Vol. 23 (5), page(s): 1157-1180.

▸ [9] **A. Bertrand**, **W. Maxwell**, **X. Vamparys**, "Do AI-Based Anti-Money Laundering (AML) Systems Violate European Fundamental Rights?," *International Data Privacy Law*, Oxford University Press, 2021.

▸ [10] F. Acquatella, **V. Fernandez**, **T. Houy**, "*Les stratégies de plateformes analysées sous le prisme de l'intelligence artificielle*," *Question(s) de Management, éditions EMS*, 2020, n°30 (4), pp.63.

# INTERACT —

## Group leaders

Françoise Détienne   Christian Licoppe
(CNRS)

*Supporting innovation
with creativity and participatory
approaches, and studying
appropriation and uses*

## Keywords
Interaction, Activity, Mobility, Collaboration, Community, Interaction design, Creative design, User experience, Technology-mediated interaction, Social robotics

## Mots-clés
Interaction, Activité, Mobilités, Collaboration, Communauté, Design d'interaction, Conception créative, Expérience utilisateur, Interaction médiée par la technologie, Robotique sociale

### See also
▶ Inclusive, innovative and reflective societies, p. 13

### Web links to the team
telecom-paris.fr/interact
### Publications
telecom-paris.fr/hal-interact

The research carried out by INTERACT is firmly embedded in innovation issues: supporting innovation with creativity and participatory approaches and the study of appropriation and uses. INTERACT relies on a combination of methods—interaction analysis, analysis from the subject's point of view, participatory methods—to make use situations visible. The group organizes its work, structured around ANR and H2020 programs, into four themes.

The **Interaction and activity** theme focuses on understanding the organization patterns of human interactions, especially in environments with a high digital technology component. Research projects study the organization of communication exchanges in "fragmented ecologies", the interactions in institutional and professional environments, where forms of speech may be constrained, and the interactions between humans and virtual agents and robots.

The **Collaboration and Communities** theme aims to understand and equip design and collaboration activities, understood as activities that are productive, targeting the co-development of a physical or digital artifact, constructive, targeting the co-development of knowledge in an interaction and socio-relational, targeting the regulation of emotions and the development of a collective approach.

The **Mobility** theme analyzes the in situ practices of technology-driven mobility: behavior linked to the new forms of mobility in urban settings, taking into account attention, risks and constraints, shared mobility, their use and appropriation, new mobile services and mobility in a transnational context, such as migrants and diasporas.

Lastly, the **Design and Creativity** theme focuses, with a non-deterministic view of creativity, on analyzing and understanding the socio-cognitive, emotional and technical processes involved in design, as well as experimenting with new situations, methods and devices for creativity.

Permanent members
Membres permanents **11**

Publications
Publications **40**

## Data physicalization and design

"Input visualizations" are visual and physical representations that are designed to collect (and represent) new data. Despite numerous examples in daily use, they have received very little attention and the benefits, trades-offs, design patterns, and even the language necessary to describe them remain unexplored. Article [3] highlights a set of recent examples and introduces vocabulary to describe them. We study the similarities and differences between the approaches carried by these examples and classical output-driven information visualization. We pose a set of provocative questions to drive discussion about the role of input visualizations, their relationship to traditional visualization designs and norms, and their usefulness, in a call to the community to help us identify additional examples of visual structures and consider their implications.

These input visualizations are a subset of physical data representations. Physical representations of data offer physical and spatial ways of looking at, navigating, and interacting with data. This was covered in the IEEE Computer Graphics and Applications Special issues on Data Physicalization [DP1, DP2]. Also called physicalizations, they allow data to be displayed through the geometric or physical properties of an artifact. However, designing and producing physicalizations require expertise in both visualization and physical fabrication. It is time-consuming, laborious, and costly. Physical rendering makes the visual presentation perceivable by bringing it into existence in the physical world: it is not often a simple, straightforward process, and it requires interdisciplinary know-how from visualization and computer graphics to design and fabrication, and human-computer interaction. The survey from research literature and physicalization practices presented in article [4] provides an overview of physicalizations, their classifications, visual representation formats, and their target datasets. It focuses on the rendering phase of the extended Infovis pipeline, in order to motivate artists and industry practitioners to "render" physicalizations using digital design and fabrication tools. Moreover, the design space [5] that will be published at the 2022 Human Factors in Computing Systems Conference will provide a summary across the different communities working in that topic, and will discuss emerging trends, and identify underlying open challenges.

## Mobility, autonomous vehicles

New mobility modes (e.g. soft modes, car-pooling, autonomous and shared vehicles) offer great promises as well as new concerns: sources of trust and mistrust, sources of comfort and discomfort, new forms of cooperative interactions and excessive connections of workers with communication technologies. To tackle these problems, it is useful to adopt an ergonomic psychology perspective, focused on the analysis of the activity in real or projected contexts and the users' actual experience (flow of action, cognition, feelings and sensations).
In the context of future shared robot taxis, article [1] shows how the guided imaginary projection methodology, which aims to have future users "experience" the service in an imaginary way, has been used as an effective way for projecting a group of people in situations—such as the absence of a driver—that cannot yet be tested in the real world. This study allowed us to identify a large number of sources of comfort and discomfort, and to draft some recommendations for the design of such services.

Article [2] explores the subjective feelings and objective risks of driving in an automatic mode when holding or not holding the steering wheel, in situations of driving on highways including critical events forcing the driver to take over. The situation awareness, feelings of safety, control, attention, on-board activities, psychological and physical comfort were analyzed on the basis of in-depth interviews right after the activity.

▸ [1] **B. Cahour**, **M. Hoarau**, and A. Rossi, "Sharing an Autonomous Taxi Without a Driver Through Guided Imaginary Projection to Identify Sources of (Dis)comfort," in *Proceedings of the 21st Congress of the International Ergonomics Association* (IEA 2021), Cham, 2021, pp. 631–638.

▸ [2] **B. Cahour**, J.F. Forzy, and A. Koustanaï, "Driving a Partially Automated Car with the Hands On or Off the Steering Wheel: Users' Subjective Experiences," in *Proceedings of the 21st Congress of the International Ergonomics Association* (IEA 2021), Cham, 2021, pp. 583–592.

▸ [3] **S. Huron** and W. Willett, "Visualizations as Data Input?," *alt.VIS, a workshop co-located with IEEE VIS* 2021. 2021.

▸ [4] H. Djavaherpour, F. Samavati, A. Mahdavi-Amiri, F. Yazdanbakhsh, **S. Huron**, R. Levy, *et al.*, "Data to Physicalization: A Survey of the Physical Rendering Process," *Computer Graphics Forum*, vol. 40, no. 3, pp. 569–598, 2021.

▸ [5] S. Bae, C. Zheng, M. E. West, E. Yi-Luen Do, **S. Huron**, and D. Albers Szafir, "Making Data Tangible: A Cross-disciplinary Design Space for Data Physicalization," in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, New York, NY, USA: Association for Computing Machinery, 2022.

▸ [DP1] T. Hogan, U. Hinrichs, **S. Huron**, J. Alexander, and Y. Jansen, "Data Physicalization," *IEEE Computer Graphics and Applications*, vol. 40, no. 6, pp. 21–24, 2020.

▸ [DP2] T. Hogan, U. Hinrichs, **S. Huron**, J. Alexander, and Y. Jansen, "Data Physicalization—Part II," *IEEE Computer Graphics and Applications*, vol. 41, no. 1, pp. 63–64, 2021.

# SID

## Group leaders

**Valérie Beaudouin**    **Christophe Prieur**

**Contact** David Bounie

The SID team is evolving in 2022. It now brings together sociologists, managers and lawyers from Télécom Paris, and becomes the Digital Technologies, Organization and Society (DTOS) team. This new team is led by David Massé and Winston Maxwell.

## Keywords
Science and technology studies, Usages, Participation, Democratization, Amateurs, Citizens, Digital labor, Design, Art-science

## Mots-clés
Sociologie des techniques, Usages, Participation, Démocratisation, Amateurs, Citoyens, Digital labor, Design, Art-science

**See also**
▸ Artificial intelligence and platform labor, p. 13

**Web links to the new team**
telecom-paris.fr/sid
**Publications**
telecom-paris.fr/hal-sid

*Transformation of social practices, new forms of collective contribution and organization, power issues and forms of exploitation which may arise*

Very rarely does an established innovation resemble the innovation its inventors had in mind initially. Observing the differences between the two—by way of field surveys on how users appropriated it and on the design process in labs—enables innovations to be adapted and to lessen any societal controversies which may have emerged. This is the context in which SID conducts its research, which is structured around several ANR (French National Research Agency) programs and addresses two key themes:

**The Internet and Society** focuses on the transformation of social practices in the context of the Internet. In particular, this theme explores digital cultures and their innovative tools for creation and mediation, how the public arena is changing, individual usage and the new forms of collective contribution and organization (communities of friends, citizens, experts and activists). It also pays particular attention to power issues and forms of exploitation which may arise (governance of digital technology, microwork, surveillance, breach of privacy and so on).

**Art and Science** (design and creation), explores how art practices can lead to new representations of society and science. Art becomes a tool for learning, understanding and experiencing current societal shifts and social debates. SID members are well aware that art contributes to the advancement of knowledge. Most of them are art practitioners themselves or involved in personal creative projects. Exploring creative processes, analyzing forms of creativity and developing teaching and research tools that draw on art and design all provide a better appreciation of the role that creativity plays in our understanding of the world and in our ability to develop thinking skills and a sense of responsibility.

SID participates in the CNRS Research Center "Internet, AI and Society."
cis.cnrs.fr/presentation-gdr

| | | | |
|---|---|---|---|
| Permanent members Membres permanents | **9** | Thesis Thèses soutenues | **2** |
| Patents Brevets | **4** | Publications Publications | **15** |

| Titre de la page | Date de création | URL | Nombre de contributeurs | Nombre de contributions |
|---|---|---|---|---|
| Pandémie de Covid-19 | 19 janv. 2020 | https://fr.wikipedia.org/wiki/Pandémie_de_covid-19 | 1471 | 8485 |
| Pandémie de Covid-19 en France | 29 fév. 2020 | https://fr.wikipedia.org/wiki/Pandémie_de_covid-19_en_France | 1041 | 7264 |
| Désinformation sur la pandémie de Covid-19 | 26 mars 2020 | https://fr.wikipedia.org/wiki/Désinformation_sur_la_pandémie_de_covid-19 | 113 | 735 |
| Test diagnostique du SARS-CoV-2 | 3 mars 2020 | https://fr.wikipedia.org/wiki/Test_diagnostique_du_SARS-CoV-2 | 71 | 1114 |
| Didier Raoult | 10 déc. 2008 | https://fr.wikipedia.org/wiki/Didier_Raoult | 381 | 2136 (1175 en 2020) |

## Social media and crisis management: a focus on the pandemic

Several events in the recent years, such as terrorist attacks or natural disasters, have confirmed the opportunities and challenges raised by social media during a crisis management, these media constituting an infrastructure for real and virtual collaborations. New practices induce a cultural and institutional change and call for a reconfiguration of the channels of information (flow and understanding) from both citizen and institutional response sides (upstream and downstream). This has been the main object of study for MACIV (Management of citizens and volunteers: the social media contribution in crisis situation), an ANR project that ended in 2021. The worldwide health crisis provided an opportunity for a novel and unanticipated situation.

Framed in the field of crisis informatics, articles [1,2] focus on the ways citizens have been collectively making sense of the Covid-19 pandemic through social media. We conducted a qualitative investigation mobilizing an ethnographic analysis of the Wikipedia pages related to the pandemic, as well as interviews with its contributors and a member of Wikimedia Foundation-France. We show how Wikipedians' know-how allows a reading of the current dramatic event, a vector of uncertainty. We also shed light on the specificities of the health crisis compared to other so-called civil security crises. Specifically, the construction of meaning related to the pandemic has required an unprecedented reorganization of how to edit and publish these pages. In particular, we took the successive stages of the pandemic and the related disinformation phenomenon into account.

▶ [1] **S. Bubendorff**, and **C. Rizza**, "*Produire collectivement du sens en temps de crise : l'utilisation de Wikipédia lors de la pandémie de COVID-19*" ("Making sense in a time of crisis: Wikipedia use during the Covid-19 pandemic"), *Communiquer*, no. 32, pp. 83-102.

▶ [2] **S. Bubendorff**, and **C. Rizza**. "*Wikipédia Face à La Crise Sanitaire.*" ("Wikipedia faced with the pandemic") *Annales Des Mines - Enjeux Numériques*, no. 14, June 2021.

## Sovereignty of the Russian Internet: surveillance and censorship via the infrastructures

The following articles contribute to the sociology of digital platforms and the study of "governance by algorithms". They provide a glimpse into a particular country that is swiftly controlling its "digital sovereignty": Russia. First and foremost, Russia is a country where US platforms did not massively dominate the search engine industry. It has therefore been possible to develop services that act nowadays as "algorithmic gatekeepers", which raise questions in terms of news diversity and possible manipulation. Article [3] studies this phenomenon through the example of the Yandex.News aggregator, a key asset in the Russian government's overall control of the country's media and digital public sphere.

Secondly, the Russian Internet has seen an increase in both legal control and centralization of technical infrastructure. In article [4], two kinds of "black boxes"—surveillance systems and traffic filtering solutions—are closely studied, and their financial, ethical and political concerns are discussed.

▶ [3] F. Daucé, and **B. Loveluck**, "Codes of Conduct for Algorithmic News Recommendation: The Yandex.News Controversy in Russia," *First Monday* 26, no. 5, May 2021.

▶ [4] K. Ermoshina, **B. Loveluck**, and F. Musiani. "A Market of Black Boxes: The Political Economy of Internet Surveillance and Censorship in Russia," *Journal of Information Technology and Politics*, April 2021.

## Dialogue of difference in controversial societies

In a situation of uncertainty, it is predictable and legitimate that science should experience controversy, where arguments in search of proof and facts are confronted with questions. Whereas controversy is useful in modern societies to describe living science, in the process of being made, it is being pushed to its utmost limits in the media cacophony, where it closes the door on any understanding of the plurality of points of view.

Article [5] is a plea for practicing a dialogue of difference, which uses the idea of interconvictionality, the art of "making the diversity of beliefs and exchanges of views the complete opposite of a confrontation between antagonistic 'truths', rather it is the source of individual and collective enrichment, the very foundation of democracy". The possible existence of such a skill is precisely what makes it possible to move from parallel beliefs to a dialogue of difference. By bringing to light all the visible signs of articulation between beliefs (such as reformulating the position of the other speaker to indicate what we understand and recognizing an area of agreement on this or that point), and by operating at various levels (social networks, media and senior officials), the dialogue of difference would re-establish links among citizens and restore a degree of appeasement in democracy.

▶ [5] **O. Fournout**, "*Ce dialogue des différences que les controverses mettent à mort.*" ("A dialogue of difference killed by controversies") *Diasporiques*, no 51, October 2020.

02

# Innovation & transfer

Innovation & transfert

# Editorial
## Édito

*The strength of the Télécom Paris incubator in supporting projects with high technological added value, such as Exotec, is that it can cut across all the key digital themes. We are able to provide strategic, financial, and technological support to address different markets and propel projects to the top of FrenchTech!*

**Yann Aprile Bouché,**
*Head of the Télécom Paris incubator*

**Nicolas Glady,** Dean and President

EN By linking together training and research with economic and social applications, Télécom Paris positions itself as a key player in digital innovation and entrepreneurship.

Over the past few years, the School has been supporting and developing the spirit of innovation in researchers, with targeted partnerships in key application fields (transport and mobility, digital health, safety and security, energy and climate, etc.), the transfer of research output to its partners and the development of innovative startups. It is mindful of protecting researchers' know-how and results by filing patents and software derived from their work. No fewer than 200 patents and software were filed in another very successful year.

Télécom Paris trains and encourages its students in designing and implementing their projects and pays particular attention to entrepreneurship in PhD students. Aware of the challenges, the School, together with the Institut Polytechnique de Paris, redouble their efforts to raise awareness of spinoff projects and encourage PhD students to engage after their course in such initiatives. Which is why projects initially conducted as part of undergraduate and graduate programs increasingly lead to the creation of startups and spinoffs.

A precursor in incubation since 1999, Télécom Paris has to date assisted more than 500 companies—including some of the best FrenchTech startups—in the acceleration and development of their products and services. Thanks to close ties with the Institut Polytechnique de Paris and its network of funders, the School's incubator, Télécom Paris Novation Center, is gaining a strong reputation in the digital field in Paris. It now boasts a unicorn, a business steeped in innovation, valued at more than a billion dollars. Exotec is the first French industrial unicorn.

*We are proud to present to you here the work of Exotec, TALia and some of the projects led in 2021, which embody our unique vision of entrepreneurship in the key fields of the digital world.*

FR En assurant le lien entre formation, recherche et valorisation économique et sociale, Télécom Paris se positionne comme un acteur clé de l'innovation et de l'entrepreneuriat dans le domaine du numérique.

Depuis plusieurs années, l'École soutient et développe l'esprit d'innovation de ses chercheurs par le développement de partenariats ciblés vers des domaines d'application clés (transport et mobilité, santé numérique, sécurité et sureté, énergie et climat...), le transfert des résultats de ses recherches vers ses partenaires, et le développement de startups innovantes. Elle veille à la protection du savoir-faire et des résultats de ses chercheurs par le dépôt de brevets et de logiciels issus de leurs travaux. Cette année encore, pas moins de 200 brevets et logiciels ont été déposés.

Télécom Paris forme et encourage ses élèves dans la conception et la réalisation de leurs projets, et porte une attention particulière à l'entrepreneuriat de ses doctorants. Conscients de l'importance de cet enjeu, l'École, et l'Institut Polytechnique de Paris, doublent d'effort pour sensibiliser et encourager les doctorants à se lancer dans des projets de spinoff à l'issue de leur formation. Il n'est ainsi pas rare que des projets initialement conduits dans le cadre des formations étudiantes ou doctorales aboutissent à la création de startups ou de spinoffs.

Précurseur de l'incubation depuis 1999, Télécom Paris a accompagné à ce jour plus de 500 entreprises, parmi lesquelles certaines des meilleures startups de la FrenchTech, dans l'accélération et le développement de leurs produits ou services. Grâce à ses liens étroits avec l'Institut Polytechnique de Paris et son réseau de financeurs, le Télécom Paris Novation Center, l'incubateur de l'École, s'affirme comme une référence dans le monde du numérique à Paris. Il peut désormais se targuer de la création d'une licorne - ces entreprises marquées par l'innovation, dont la valorisation dépasse le milliard de dollars - Exotec, qui est la première licorne industrielle française.

*C'est donc avec fierté que nous vous parlons dans les pages qui suivent d'Exotec, de TALia et de quelques-uns des projets de 2021 qui ont incarné notre vision unique en matière d'entrepreneuriat dans les domaines clés du numérique.*

# Deep Tech Factory

Télécom Paris is actively engaged in supporting innovation via a number of teaching and research projects in partnership with industry. Projects provides students and research teams with practical usage cases and enjoys world-class multidisciplinary expertise. Cooperation varies widely in form and involves many faculty members in different scenarios, including partnership-based PhD (CIFRE Program), bilateral contracts, ANR- and H2020-type projects, joint labs, teaching and research chairs, technology transfer and the support of startups and spinoffs. In 2019, Télécom Paris derived income from the commercialization of a dozen patents and software programs.

Among the current teaching and research chairs at Télécom Paris, which were created in partnership with companies and in some cases with the support of the Fondation Mines Télécom (see pages 10-11), the Innovation & Regulation in Digital Services Chair is particularly devoted to the field of digital innovation, where it contributes to explore and develop a landmark theoretical framework on two main themes: the creation, production, development and consumption of immaterial services; the dynamics of innovation, the principles and tools of regulation and their economic effects on medium and long terms.

▸ The Chair organized its 12th international Conference on Digital Economics on April 7-9, 2021, with Digital Economics Paris, a network of researchers based in and around Paris. Platforms and pricing, reputation, data, privacy, competition and regulation, and antitrust in the digital economy were some of the topics discussed during the sessions.

▸ The 7th edition of the Digital Economics Summer School of AFREN (the Francophone Association for Research in Digital Economics), with the support of the Innovation & Regulation in Digital Services Chair, took place in Télécom Paris's new campus, June 28-29, 2021.

▸ https://innovation-regulation.telecom-paris.fr

In 2021, there were 54 ongoing PhDs in the form of Industrial Agreements for Training through Research (CIFRE), in partnership with companies such as Nokia, Orange or EDF. Since 2013, Télécom Paris has also been offering a one-semester training program on innovation: Research and Innovation Master Projects (PRIM). These provide an opportunity to develop a real innovation project or the creation of a business, liaising with a Télécom Paris lab and/or a corporate partner. Patents are the natural products of research. As such, they attract particular attention from the Ministry for the Economy, as they aim to serve national and European markets and represent an increasing demand from industry partners who would like to protect their work before it is disseminated.

## Byzantine Fault Tolerant systems with a better latency than the state of the art

Obtaining consensus in large blockchains based on Byzantine Fault Tolerance (BFT) systems, such as Diem (Facebook) or Ethereum 2.0, needs a seven message exchanges under favorable circumstances. The state of the art, Hotstuff, achieves linear communication complexity, but at the cost of additional latency, due to one more round-trip with the leader. Matthieu Rambaud (Associate professor in the newly created C² team) and its colleagues from CWI Amsterdam, Netherlands, reduced the time to five message exchanges. The patent pending invention combines the use of digital signatures and cryptographic zero-knowledge proofs.

▸ M. Abspoel, T. Attema, and **M. Rambaud**, "Brief Announcement: Malicious Security Comes for Free in Consensus with Leaders," in *Proceedings of the 2021 ACM Symposium on Principles of Distributed Computing*, New York, NY, USA, 2021, pp. 195–198.

Patents invented in 2020-2021
Brevets inventés en 2020-2021
**28**

Software filed in 2020-2021
Logiciels déposés en 2020-2021
**8**

# Télécom Paris
## Novation Center



Seen as one of the leading incubators for digital sciences innovation in France, the incubator based at the heart of Télécom Paris turned twenty-two in the spring of 2021. It has led to the creation of more than 500 companies. With a satellite at Sophia Antipolis since 2006, it became known as ParisTech Entrepreneurs in 2009, serving students, researchers and alumni from ParisTech schools. It is also open to entrepreneurs in Paris from different backgrounds. In 2016, SME Builder—a startup accelerator—was set up to complement the other facilities. Its aim is to create strong SMEs in the digital industry. The Paris site hosts 25 startups at any given time for a period of 18 months. Companies with disruptive projects based on digital technologies and usage are selected for their potential synergy with Télécom Paris' research labs and their potential for business, growth and international development. They benefit from a mentoring program and partnerships with many players in business and science.

**On December 9, 2019, the incubator changed its name to Télécom Paris Novation Center.**

**Web**
incubateur-telecomparis.fr

## MIMOPT

MIMOPT Technology is a spinoff created in April 2021 by Ghaya Rekaya (Professor, ComNum team), Yves Jaouen (Professor, GTO team), and Akram Abouseif, their former PhD student. MIMOPT develops and markets innovative digital signal processing solutions to maximize data throughput of optical fiber communication systems. Its novel, patented solutions enable high order wave length division multiplexing and spatial division multiplexing on the fiber, while reducing interference between parallel channels.

This spinoff is the fruit of more than thirteen years of research work at Télécom which has resulted in a patent portfolio consisting of fourteen patent clusters.

MIMOPT Solutions guarantee the best performance for already deployed optical fibers such as standard single-mode fiber with single wavelength or Wavelength-division-Multiplexing; as well as for future technologies such as Space-division-multiplexing that can be addressed by multi-core fiber or multi-mode fiber. These solutions can be used for several applications: in ultra-long-haul transmission in submarine, in long-transmission in core networks and metro networks, and in access networks and data center. The objective is to reach the maximum capacity limit with the lowest complexity.

In addition, MIMOPT is working on multi-mode and multi-core fibers, which will be the fibers of tomorrow. These fibers will solve throughput problems, provided that they do not generate prohibitive additional energy costs. The team is proposing to address this issue by developing coding solutions that will save the power injected into the fibers.

▸ http://mimopt.com

## 1999-2021 data / chiffres

**+500**
Startups created
Startups créées

**200**
Patents filed
Brevets déposés
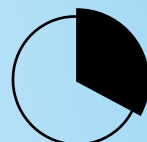
**33%**
... of startups have at least one founder with a PhD from a university or *grande école*
... des startups ont à minima un docteur universitaire ou d'une grande école parmi leurs fondateurs

**1080 M€**
...raised since 2015 by startups in or having left the incubator
...de fonds levés depuis 2015 pour les startups issues de l'incubateur et en cours

**+5000**
Jobs created
Emplois créés

# Startups and Spinoffs
## hitting the headlines

# 17 January 2022

## Exotec

In early 2022, Exotec, a leader in global warehouse robotics, announced raising $335M in Series D funding at a $2B valuation, becoming France's first industrial unicorn.

Now based in Lille, France, with engineers and technicians all over the world, the company grew out of the Télécom Paris incubator in 2015. Renaud Heitz and Romain Moulin, co-founders of Exotec, had then decided to design a robotic order-picking system with a fleet of robots—the Skypod—capable of moving in three dimensions. "*The Télécom Paris incubator enabled us to benefit immediately from a structure to start working, as well as top quality training, some of the precepts of which are still useful to us in the life of Exotec,*" says Romain Moulin, CEO, who adds "*It also enabled us to facilitate the first fundraising round by putting us in contact with its network. Finally, and this is probably one of the most important elements in our opinion, we were accompanied by a passionate and caring team.*"

▸ https://www.exotec.com/en/fundraising-2022

## Secure-IC

In March 2021, Secure-IC announced partnership with U.S. DARPA to foster security technology innovation; as part of the DARPA Toolbox initiative. The spinoff created in 2010 through Télécom Paris laboratories, with three of the school's researchers as co-founders, Jean-Luc Danger, Sylvain Guilley and Laurent Sauvage, provides security solutions for connected objects and embedded electronic systems (mobile phones, electronic passports, credit cards, automobile electronics etc.) in order to combat cybercrime.

The SSH team and Secure-IC continued in 2021 their joint research on side-channel attacks.

### Protecting against fault injection and auxiliary channel attack via code-based masking

Masking is one of the best-studied countermeasures to protect cryptographic implementations against side-channel attacks due to the favorable provable security it provides. The core idea underlying any masking scheme is to split the sensitive (key-dependent) variables into several shares and perform independent computations on masked variables only. The generalized code-based masking (GCM) is the most generic scheme in this respect. Polynomial masking is a special case of GCM, which is built upon Shamir's secret sharing (SSS) scheme.

In article [1], we propose a unified analytic approach to quantify information leakage in generalized code-based masking (GCM), and show how to select optimal codes for GCM. In articles [2,3,4], we apply our leakage quantification approach to Inner Product Masking (IPM) and present optimal linear codes in several cases. At last, as presented in thesis [5], the unified leakage quantification approach is fully verified by empirical attack-based evaluations. Moreover, an interesting extension of IPM for fault detection is proposed in article [6], which can be deployed to mitigate both side-channel analysis and fault injection attacks simultaneously.

▸ [1] **W. Cheng**, **S. Guilley**, C. Carlet, **J.-L. Danger**, and **S. Mesnager**, "Information Leakages in Code-based Masking: A Unified Quantification Approach," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2021, no. 3, pp. 465–495, Jul. 2021.

▸ [2] **W. Cheng**, **S. Guilley**, C. Carlet, **S. Mesnager**, and **J.-L. Danger**, "Optimizing Inner Product Masking Scheme by a Coding Theory Approach," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 220–235, 2021.

▸ [3] **W. Cheng**, **S. Guilley**, and **J.-L. Danger**, "Categorizing all linear codes of IPM over $\mathbb{F}_{2^8}$," *Cryptography and Communications*, vol. 13, no. 4, pp. 527–542, Jul. 2021.

▸ [4] **W. Cheng**, **Y. Liu**, **S. Guilley**, and **O. Rioul**, "Towards Finding Best Linear Codes for Side-Channel Protections," *10th International Workshop on Security Proofs for Embedded Systems* (PROOFS'2021), Beijing, China, Sep. 2021.

▸ [5] **W. Cheng**. "What Can Information Guess? Towards Information Leakage Quantification in Side-Channel Analysis." PhD Thesis, Télécom Paris, Institut Polytechnique de Paris, NNT: 2021IPPAT044, 2021.

▸ [6] **W. Cheng**, C. Carlet, **K. Goli**, **J.-L. Danger**, and **S. Guilley**, "Detecting faults in inner product masking scheme," *Journal of Cryptographic Engineering*, vol. 11, no. 2, pp. 119–133, Jun. 2021.

▸ https://www.secure-ic.com

# Joint laboratories

"Alternative" joint laboratories, typically set up for a five-year period, are groupings that bring together the research potential of Télécom Paris, together with that of one or several companies, higher education institutions and research centers on a given topic.

## Internet of Things and trusted telecommunications for industry, energy services and customers

SEIDO Lab was created in 2012 by EDF R&D and Télécom Paris to conduct joint research on the Internet of Things and cybersecurity applied to EDF's industrial challenges. SEIDO Lab has now been extended to new partners, namely CNRS LAAS, Télécom SudParis, and recently École des Mines de Saint-Étienne.

2022 marks the launch of a new five-year period for SEIDO Lab. The new research program is positioned on IoT and trusted telecommunications for energy services. It aims to ensure the trust, efficiency and security of connected objects and telecommunications—including 5G and blockchain—and of the services supported by these technologies. Application domains range from smart homes and smart buildings (in particular via the combination of trusted energy and other services provided by IoT), smart grids and energy communities (for secure, decentralized and multi-scale energy management, including local production, storage, mobility), and the management of cyberattacks in industrial environments.

The laboratory addresses cross-cutting issues such as monitoring hardware and software vulnerabilities and intrusions, decision-making for risk management and the deployment of security mechanisms, user trust in IoT system decisions, compliance assurance of security policies or services, deployment of flexible, adaptive, multi-scale, and interoperable services provided by connected objects.

SEIDO Lab organizes an annual workshop, an opportunity to present the results of the work carried out by its researchers and to invite keynotes on SEIDO Lab's core topics. To date, SEIDO Lab has filed more than ten French and European patents, published more than fifty peer-reviewed articles and papers in international journals and symposiums, in addition to four book chapters. So far, six PhD theses have been defended, five are ongoing, and five more are to start in 2022. ACES (p. 44), DIG (p. 50), RMS (p. 58) and SSH (p. 42) contribute to the work of SEIDO Lab on behalf of Télécom Paris.

▸ https://seido-lab.com

## Other joint laboratories

Télécom Paris' strategic research lines extend to other joint laboratories: ISA (Security and Identity), Exalt Design Lab (Design in Companies), BibliLab (Digital Heritage of Libraries), Lincs (Internet of the Future), and ΣΚΦ, the first joint lab involving different members of Institut Polytechnique de Paris, created by ACES at Télécom Paris, COSYNUS (École Polytechnique) and U2IS (ENSTA Paris).

AllegoRI, the joint research activity in 3D science and technology between Télécom ParisTech and the Allegorithmic SME (acquired by Adobe in 2019) resulted in a technology transfer agreement.

## TALia, natural language processing (NEW)

Télécom Paris and onepoint have joined forces in 2021 to create TALia, a joint research laboratory dedicated to natural language processing. This laboratory will focus its work on three issues: to humanize our relationships with machines; to maintain linguistic sovereignty, via the control over the data sets used to train automatic language processing models, and the defense of the French language; to contribute to the transmission of contemporary knowledge, the empowerment of society and the empowerment and training of its individuals.

TALia will conduct work on the interaction between machines and humans with a focus on creating solutions capable of understanding natural language integrated in applications found in our daily lives: speech synthesis, spell checker, word prediction on smartphones, and the analysis of opinions and emotions.

The laboratory, held by Tiphaine Viard and Thomas Bonald (DIG), has programmed its work over a five-year cycle and aims to support four to five theses.

## Valeo.ai

Valeo.ai is the first artificial intelligence research center for automotive applications, started in 2017 with the CEA, Inria, Mines ParisTech and Valeo as partners. In 2021, we contributed to the release of the Carrada dataset, the first public automotive radar dataset with cars, cyclists and pedestrians with precised range-angle-doppler annotations in the raw signals.

▸ https://www.valeo.com/en/valeo-ai

# Open Source Software

## Logiciels libres

**EN** Digital science could not function without software to design, test, model, assess and apply its research activities. Télécom Paris has a specific policy of advocating free open source software, which forms an integral part of its innovation strategy, along with its academic and industrial partners. Télécom Paris COSI (Center for Open Software Innovation) groups all the projects staff have initiated or contributed to. Some thirty software packages have been made available to everyone. They cover four major fields: Multimedia, Data science and artificial intelligence, Programming and Smart networks and objects. We present here a selection of new and recently updated projects.

**FR** Les sciences du numérique ne sauraient se passer de logiciels pour concevoir, tester, modéliser, évaluer ou mettre en application les travaux de recherche. Télécom Paris mène en particulier une politique volontariste pour les logiciels libres qui sont une partie intégrante de sa stratégie d'innovation, menée avec ses partenaires académiques et industriels. L'ensemble des projets auxquels contribuent des personnels de l'école ou bien qu'ils ont initiés est réuni au sein du Centre pour l'innovation des logiciels libres (COSI) de Télécom Paris. Une trentaine de logiciels sont ainsi disponibles pour tous, couvrant quatre grands domaines : Multimédia, Science des données et intelligence artificielle, Programmation, Réseaux et objets communicants. Nous présentons ici une sélection de projets nouveaux ou qui ont connu récemment des mises à jour majeures.

**Web**
telecom-paris.fr/cosi

### Software Heritage, the largest source code archive in the world

Software Heritage is an initiative launched in 2015 by Roberto Di Cosmo and Stefano Zacchiroli (the latter joined the ACES team in September 2021, see page 19). It was publicly announced in 2016 with the support of Inria and Unesco, and the initial and ambitious goal of archiving all available software in source code format and its development history. Five years later, Software Heritage contains more than ten billion source files collected and more than two billion commits from over 160 million software development projects.

From an industry perspective, Software Heritage provides an open knowledge base that helps with the traceability of open source software, which is increasingly present in most commercially available hardware devices (such as IoT). From an academic point of view, it opens up new avenues of research in code quality improvement, vulnerability scanning, software security, and developer productivity.

## River: learning from data streams in Python

The DIG team is part of the developer team—a large community of practitioners and researchers—of **river**, a Python (and Cython) package for dynamic data streams and continual learning. Merged from the two most popular packages for stream learning in Python, Creme and scikit-multiflow, its ambition is to be the go-to library for undertaking machine learning on streaming data.

Indeed, the conventional approach in machine learning is to process data in batches or chunks. This has strong practical consequences on the way to store data and to train models. A different approach studied by DIG members for several years is to treat data as a stream: data is not stored and models continuously learn one data sample at a time. River has thus been developed to satisfy these evolving needs of a major machine learning community.

River provides multiple state-of-the-art learning methods, data generators/transformers, performance metrics and evaluators for different stream learning problems. Supported applications are as diverse as those found in traditional batch settings, aimed at industrial applications as well as academic research, including: classification, regression, clustering and representation learning, multi-label and multi-output learning, forecasting, and anomaly detection.

https://riverml.xyz/latest

▸ J. Montiel, M. Halford, S. M. Mastelini, G. Bolmier, R. Sourty, R. Vaysse, A. Zouitine, H. M. Gomes, J. Read, **T. Abdessalem**, **A. Bifet**, "River: Machine Learning for Streaming Data in Python." 2020.

Among our major contributions, see also: **GPAC**, a platform developed under the supervision of Jean Le Feuvre (MM), covering various aspects of multimedia, from formats such as MP4 or MPEG-2 TS to presentation technologies; **RAMSES**, a model transformation and code generation tool for the design of critical embedded real-time systems (ACES); **TTool**, a toolkit for editing UML and SysML diagrams and the simulation and formal verification (safety, security, performance) of those diagrams (LabSoC).

## YAGO 4 a "reason-able" knowledge base

In its simplest variant, a knowledge base (KB) takes the form of a labeled graph, where the nodes are entities and the edges represent the links between these entities in the real world. **Knowledge graphs** have been constructed automatically from web contents and text sources, and they have become a key asset for AI applications such as personal assistants, text analysis, and search engines. In particular, such knowledge can be harnessed to interpret textual phrases semantically in news, social media and web tables, and contributes to question answering, natural language processing and data analytics.

In an extended abstract [1], Fabian M. Suchanek (Professor, DIG team) outlines the challenges in the construction, completion, querying and maintenance of KBs and lists application domains such as combinatorial creativity and digital humanities. Indeed, research has made extraordinary progress in the automated construction of KBs in recent years.

YAGO was one of the first academic projects to build a knowledge base automatically. YAGO 4, released in September 2020, aims to combine the best of two worlds (article [2]). It collects the facts about instances from Wikidata, but it forces them into a rigorous hierarchy with semantic constraints. The complex taxonomy of Wikidata is replaced by the simpler and clean taxonomy of schema.org. The result is a new knowledge base, which is not just large, but also logically consistent, so that OWL-based reasoning is feasible.

YAGO 4 is a "reason-able" knowledge base. This opens a path to knowledge harvesting and complex reasoning formalisms combining, as an extension of KBs (see Extensions of knowledge bases, page 51).

http://yago-knowledge.org

For a broader view on these topics, [3] surveys fundamental concepts and practical methods for creating and curating large knowledge bases. It covers models and methods for discovering and canonicalizing entities and their semantic types and organizing them into clean taxonomies. On top of this, it discusses the automatic extraction of entity-centric properties. To support the long-term life-cycle and the quality assurance of machine knowledge, it presents methods for constructing open schemas and for knowledge curation. Case studies on academic projects and industrial knowledge graphs complement the survey of concepts and methods.

▸ [1] **F. Suchanek**, "A Hitchhiker's Guide to Ontology," 2021. DESIRES 2021 – *2nd International Conference on Design of Experimental Search & Information REtrieval Systems*, September 15–18, 2021, Padua, Italy.

▸ [2] **T. Pellissier Tanon**, G. Weikum, and **F. Suchanek**, "YAGO 4: A Reason-able Knowledge Base," in *The Semantic Web*, Cham, 2020, pp. 583–596.

▸ [3] G. Weikum, L. Dong, S. Razniewski, and **F. Suchanek**, "Machine Knowledge: Creation and Curation of Comprehensive Knowledge Bases," *CoRR*, vol. Abs/2009.11564, 2020.

# Research and innovation ecosystems

- ▸ École Doctorale de Mathématiques Hadamard
- ▸ LMH : Labex Mathématique Hadamard
- ▸ Labex DigiCosme
- ▸ Quantum, Centre de Sciences et Technologies Quantiques

## Institut Polytechnique de Paris

- ▸ **Télécom Paris**
- ▸ **École Doctorale Institut Polytechnique de Paris**
- ▸ **Centres Interdisciplinaires** (E4C, CIEDS, Hi! PARIS)

## Université Paris-Saclay

- ▸ **Maison des Sciences de l'homme Paris-Saclay**

## HEC

- ▸ FMJH : Fondation Mathématique Jacques Hadamard
- ▸ Fondation Mines-Télécom
- ▸ Institut DataIA
- ▸ SATT Paris-Saclay
- ▸ IRT SystemX
- ▸ NanoInnov

Plateau Orsay – Palaiseau – Saclay

Domaines
d'Intérêt Majeur
▸ DIM RFSI
▸ DIM MathInnov
▸ DIM SIRTEC

Systematic

Cap Digital

CEA

Inria

CNRS

IMT

Institut Carnot Télécom
et Société numérique

## 12 February 2021

### CONTINUUM Equipex

**DIVA** is involved in the **CONTINUUM** Equipex, in collaboration with the **INTERACT** team. This project will create a collaborative research infrastructure of 30 platforms located throughout France, to advance interdisciplinary research based on interaction between computer science and human and social sciences. CONTINUUM involves 37 research teams that will develop cutting-edge research programs focusing on visualization, immersion, interaction and collaboration, as well as on human perception, cognition and behavior in virtual/augmented reality, with potential impact on societal issues. It will enable a paradigm shift in the way we perceive, interact, and collaborate with complex digital data and digital worlds by putting humans at the center of the data processing workflows.



## 8/9 October 2021

### Fête de la science

200 researchers and doctoral students from IP Paris welcomed more than 1,000 visitors to discover science through creative activities.

# IP Paris
## activities
## Les activités IP Paris

Télécom Paris is co-founder of the Institut Polytechnique de Paris (IP Paris), a world-class science and technology institute created in 2019 with the École Polytechnique, ENSTA Paris, ENSAE Paris and Télécom SudParis.

Research at IP Paris is built on the recognized expertise of its founding schools in ten major disciplines, which are as many departments: Biology; Chemistry and Processes; Economics; Information, Communications and Electronics; Computer Science, Data, and AI; Mathematics; Mechanics and Energetics; Physics; Humanities, Arts, Literature and Languages; Social Sciences and Management.

Bruno Thédrez, head of the Communications and Electronics Department of Télécom Paris, is also head of the ICE (Information, Communications, Electronics) Department of IP Paris.

▸ www.ip-paris.fr

The doctoral research conducted within IP Paris is based on 30 laboratories and takes place in a high-quality scientific environment. PhD students are offered the opportunity to conduct their PhD in cutting-edge research laboratories and with companies with renowned R&D departments.

## ICE webinars

Disciplinary research at IP Paris is based on the expertise of its founding schools. Information Communications and Electronics is one of ten disciplinary research fields bringing together top researchers and PhD students at state-of-the art laboratories and facilities.

The topics of the ICE webinars in 2021 reflect the richness of the research works: Analog hardware for energy efficient AI; Indoor positioning technologies: limitless creativity to model the complexity of cities and human gaits; Structured clique networks as efficient associative memories; Trends and challenges of low-power autonomous devices; More than Moore analog electronics for harsh environment applications; Quantum sensors based on NV center in diamonds; Secure communications in quantum networks; Open software for open hardware and the licenses.

## 19 November 2020

### IP Paris Electrical Engineering Artificial Intelligence Day

During this one-day event introduced by Bruno Thedrez, head of the ICE department at IP Paris, the current research in Artificial Intelligence and Electrical Engineering has been reviewed on the following topics: AI and Information Theory; AI and Communications; AI in Electronics and Optics.
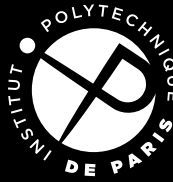All presentations are available online.

## 10 June 2021

### PhD students' day

Computer Science, Data and Artificial Intelligence (IDIA) department of IP Paris held a special day for its PhD students. Giammarco La Barbera, Léo Laugier, Yibo Quan, Raphael Achddou, Arthur Ouaknine, Nan Li, Emanuele Dalsasso, Étienne Houzé and Kamélia Daudel from Télécom Paris, presented their work during one of the three poster sessions. Lou Salaün, awarded in 2021 of the first PhD thesis prize of the IDIA Department and Pierre Laforgue, 2nd Best Thesis, presented their doctoral work in a special session.
Posters, videos, slides, articles are available at https://idia-phd-2021.ebfe.fr

IP Paris develops an ambitious innovation policy in order to increase its societal impact. Télécom Paris Novation Center is one of the three incubators of IP Paris, supporting entrepreneurs in their economic, societal and environmental innovation projects.

# IP Paris
# call for innovative projects 2021

The third call for projects to detect and support innovative projects from IP Paris laboratories ended in December 2021 with the selection of 11 projects, for a total of 840 000 euros in funding. This will help the project holders to achieve a first proof of concept and to confront their ideas with market expectations and needs.

Three projects from Télécom Paris have been selected.

---

**Aymeric de Javel**
PhD student, RMS team

**Convergent analysis of 5G networks**

**Optimization of 5G networks based on the analysis of big data from the network**

This project aims at analyzing the huge amount of available data on 5G networks using big data processing algorithms. This will address different business needs such as dimensioning, anomaly detection, troubleshooting and quality of service. In the context of network softwarization, which places automation, orchestration and telemetry at the center of operators' issues, the project focuses on the specificities of 5G: new use cases thanks to slicing and new deployment options thanks to Open RAN (Open Radio Access Networks).

The project is twofold. First, a probe able of extracting data from a 5G network is developed and a set of algorithms retrieving information from the network is implemented. In a second phase, data from the underlying IP networks are integrated in order to provide a unified RAN/IP view to operators and customers.

---

**Sumanta Chaudhuri**
Associate professor, SSH team

**ApproxiNet**

**Approximate and sparse neural network accelerator for edge AI**

This project is set in the context of a growing demand for a customized computer architecture at the edge, for deep learning applications where the AI algorithms are executed close to the source of data/sensor without sending it to the clouds/HPC servers. This exhibits various constraints such as small form factor, low power consumption and low latency, and needs fine optimization between power, speed, and accuracy. The approximate computing paradigm, which prescribes trading off accuracy for saving energy, will be explored to achieve this goal.

The project team envisions that the Edge AI market will be ultimately dominated by the ASIC solution: it gives the best possible performance achievable, and as Edge AI promises to be a very high volume market, it justifies the cost of ASIC development. Starting from a working prototype on FPGA, their short term goal is to fabricate the same design on Silicon, in 28nm TSMC process. In the long run, they will provide an Edge AI accelerator IP to OEMs, SoC manufacturers and foundries.
See also Sumanta Chaudhuri presentation during the IP Paris Electrical Engineering Artificial Intelligence Day, youtu.be/_ghYUwiyqlg&t=3966s

---

**Elsa Angelini**
Associate professor, IMAGES team

**Midway**

**Medical image harmonization by Midway**

Midway-BrainTumor is a software for processing MRI images of the brain for the calculation of difference maps and the detection of tumor growth by statistical thresholding between two examinations taken several months apart. The underlying patented technology addresses a problem that is still open, and increasingly limiting with the rise of AI based on the integration of multi-center cohorts with large variability between images from different scanners and hospitals.

The project is twofold. Further use cases will be explored in a first phase: new types of pathologies (metastasis, inflammatory, chronic), organs (in addition to brain, lungs) and modalities (in addition to MRI, CT scan). A new business model based on service provision will be proposed to health professionals and industrials on the domain.

# Carnot Télécom & Société numérique

**INSTITUT CARNOT**
**Télécom & Société numérique**

The Carnot label was created in 2006 to promote research projects undertaken by both public research players and those from the socio-economic world. Télécom Paris is a member of Carnot Télécom & Société Numérique (TSN), the first Carnot dedicated to information and communication science and technology, which is part of a network of 29 Carnot Institutes. The Carnot TSN comprises 30 research laboratories, including the LTCI at Télécom Paris, totaling more than 1,700 researchers and PhD students. It provides cutting-edge research and integrated solutions to issues related to ICT (information and communication technology).

▸ carnot-tsn.fr

## SECVD (Sécurité des véhicules de demain)

As more and more connected or autonomous vehicles travel on road networks, there are increasing opportunities for attacks on these networks, such as identity disclosure, Sybil attack, denial of service, and impersonation.

In this context, Mounira Msahli (associate professor in the ACES team) has been granted a funding by the Institut Carnot for her project SECVD (Security of tomorrow's vehicles). This project aims at formally verifying the security (authentication, confidentiality, integrity and privacy) of the Cooperative Intelligent Transport System protocols, more precisely the PKI protocol. New protocols are likely to be born following these work, and patent applications to be filed.

**See also**
▸ ACES expertise on these topics: "Secure and safe systems", page 45.

# Contact

▸ **About this document**
com.scientifique@telecom-paris.fr

▸ **Direct contacts**
firstname.lastname@telecom-paris.fr

# Credits

▸ **Writing**
Aymeric Poulain Maubant, Nereÿs

▸ **Project team at Télécom Paris**
Talel Abdessalem, Ingrid Colleau, Laurence Monéger, Stéphane Potelle

▸ **Graphique Design** ArtFeelsGood

▸ **Photo credits**
Mario Caruso on Unsplash (p. 38), CDC on Unsplash (p. 13), Michel Desnoues, Télécom Paris (pp. 3, 9, 16, 21, 22, 23, 28, 30, 31, 32, 33, 42, 43, 52, 53, 55, 63, 81, 82, 83), Fondation Mines-Télécom (p. 9), Alina Grubnyak on Unsplash (pp. 8, 12), Ali O'Connor, EXOTEC ENT (pp. 80, 84), Ghaya Rekaya (p. 83), Sebastian Scholz (Nuki) on Unsplash (p. 9), Shutterstock (pp. 7, 10,11, 22, 23, 27, 41, 44, 49, 60, 68, 71, 73, 75, 87, 93).

▸ **Printing**
ADM - 44 rue Eugène Carrière - 75018 Paris

# Special thanks to our partner companies and organizations who supported and collaborated on our research projects

**Nous remercions tout particulièrement nos entreprises et organisations partenaires qui ont soutenu et collaboré à nos projets de recherche**

- 3-5 LAB
- Accenture
- ACPR - Banque de France
- Afnor
- AG2R La Mondiale
- Agence de l'innovation de défense
- AI Builders
- Air Liquide Medical System
- Airbus Defence & Space
- Akur8
- Allegorithmic
- ALMERYS
- Amarisoft
- Amossys
- ANFR
- Arcep
- ATEME S.A.
- Atos
- Axione
- Axys
- be ys
- BnF
- BNP Paribas
- Bouygues Telecom
- Caisse des Dépôts
- Capgemini
- Cardiologs
- Carrefour
- CGI
- Cisco
- CNIL
- Converteo

- Crédit Agricole
- CRITEO
- Danone
- Dassault Aviation
- Dassault Systèmes
- Datadog
- Dataiku
- DCNS
- DEEZER
- Delta Dore
- Devoteam
- DGA
- DxOMark
- EASYRECRUE
- ECA Group
- EDF
- Ekinops
- EMAKINA
- Engie
- Eutelsat
- EXFO OPTICS
- Fabernovel
- Fondation d'entreprise Carrefour
- Gemalto
- General Electrics Healthcare
- Google
- Groupe BPCE
- Groupement des Cartes Bancaires
- Guerbet
- Headmind Partners
- HN Services
- Huawei

- IBM
- IDAaas
- Idemia
- Ikos
- IN Group
- Incepto
- Innovation Defense Lab
- InProcess
- Intelcia
- Interdigital
- Kardham Digital
- KEEN EYE
- Kering
- La Banque Postale
- La Poste
- LinkByNet
- L'Oréal
- MAIF
- Mazars
- MBDA
- MC2I
- Facebook / Meta
- Ministère des Armées
- Mitsubishi
- MVG
- Naval Group
- Navigacom
- Nokia
- Nokia Bell Labs
- Onepoint
- Orange
- OTIS

- PHILIPS FRANCE
- PMP
- PWC
- REGIENOV
- Renault
- Rexel
- Safran
- Schlumberger
- SECURE IC
- Société Générale
- Softbank Robotics
- Sonos
- Sony CSL
- Sopra Steria
- STMicroelectronics
- STREAMROOT
- Symocore
- Talan
- TDF
- TECHNICOLOR et INTERDIGITAL
- Thales
- Tiime
- TotalEnergies
- Ubisoft
- Valeo
- Véolia
- Vinci
- Wavestone
- Zaion

**Télécom Paris**
**2021**
**Review**
Rétrospective

▶ **telecom-paris.fr**