

2025 Activity Report

RESEARCH CENTRE: Inria Lyon Centre

IN PARTNERSHIP WITH: CNRS, Université Jean Monnet Saint-Etienne

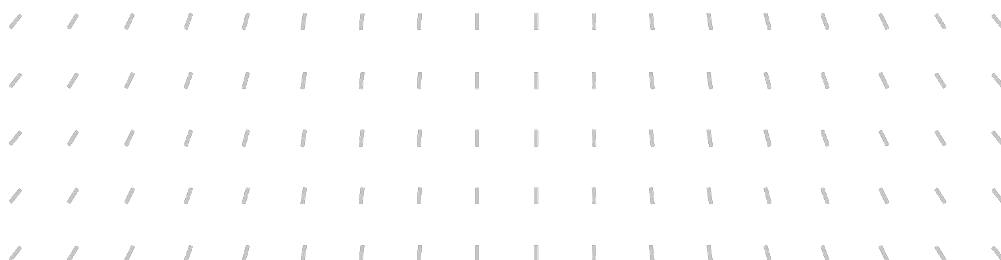

Project-Team

MALICE

MACHINE Learning with Integration of surfaCe
Engineering knowledge: Theory and Algorithms



In collaboration with Laboratoire Hubert Curien (LabHC)



Project-Team MALICE

Creation of the Project-Team: 2023 December 01

Each year, Inria research teams publish an Activity Report presenting their work and results over the reporting period. These reports follow a common structure, with some optional sections depending on the specific team. They typically begin by outlining the overall objectives and research programme, including the main research themes, goals, and methodological approaches. They also describe the application domains targeted by the team, highlighting the scientific or societal contexts in which their work is situated. The reports then present the highlights of the year, covering major scientific achievements, software developments, or teaching contributions. When relevant, they include sections on software, platforms, and open data, detailing the tools developed and how they are shared. A substantial part is dedicated to new results, where scientific contributions are described in detail, often with subsections specifying participants and associated keywords. Finally, the Activity Report addresses funding, contracts, partnerships, and collaborations at various levels, from industrial agreements to international cooperations. It also covers dissemination and teaching activities, such as participation in scientific events, outreach, and supervision. The document concludes with a presentation of scientific production, including major publications and those produced during the year.

Keywords

Computer sciences and digital sciences

- A3. – Data and knowledge
- A3.4. – Machine learning and statistics
- A5.9.1. – Sampling, acquisition
- A5.9.4. – Signal processing over graphs
- A5.9.5. – Sparsity-aware processing
- A5.9.6. – Optimization tools
- A6.3.1. – Inverse problems
- A6.3.5. – Uncertainty Quantification
- A6.5. – Mathematical modeling for physical sciences
- A8.2. – Optimization
- A8.12. – Optimal transport
- A9.2. – Machine learning
- A9.2.1. – Supervised learning
- A9.2.4. – Optimization and learning
- A9.2.6. – Neural networks
- A9.2.7. – Kernel methods
- A9.2.8. – Deep learning
- A9.3. – Signal processing
- A9.11. – Generative AI

Other research topics and application domains

- B2.6. – Biological and medical imaging
- B9.5.1. – Computer science
- B9.5.2. – Mathematics
- B9.5.3. – Physics
- B9.5.6. – Data science

Contents

Project-Team MALICE	1
1 Team members, visitors, external collaborators	5
2 Overall objectives	6
3 Research program	6
4 Application domains	8
5 Social and environmental responsibility	8
6 Highlights of the year	8
6.1 Papers	8
6.2 Inria Associate Teams	9
6.3 Nomination	9
6.4 Awards	9
7 Latest software developments, platforms, open data	9
7.1 Latest software developments	9
7.1.1 SwiftHohenbergPseudoSpectral	9
7.1.2 GUAP	9
7.1.3 GraSPy	10
7.1.4 Scikit-SpLearn	10
7.1.5 SoundScapeExplorer	10
8 New results	11
8.1 Generative AI - Flow Matching	11
8.1.1 On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity (NeurIPS'25) [2]	11
8.1.2 A Visual Dive into Conditional Flow Matching (ICLR Blogposts'25) [4]	11
8.2 Physics-informed Machine Learning (PIML)	12
8.2.1 A Bregman Proximal Viewpoint on Neural Operators (ICML'25) [6]	12
8.2.2 Provably Accurate Adaptive Sampling for Collocation Points in Physics-informed Neural Networks (ECML'25) [3]	12
8.2.3 Conformal Online Learning of Deep Koopman Linear Embeddings (NeurIPS'25) [5]	12
8.2.4 Using Signatures and Koopman Operator to Learn Nonlinear Dynamics (GSI'25)	13
8.2.5 Contextual Hypernetwork for Adaptive Prediction of Laser-Induced Colors on Quasi-Random Plasmonic Metasurfaces (ECML'25) [25]	13
8.2.6 Physics-Informed Machine Learning for Modeling CO2 Capture from Scarce Data (ICTAI'25) [18]	13
8.2.7 Image quality metrics for restricted gamut images produced by laser-induced printing on plasmonic films (Journal of the European Optical Society 2025) [9]	14
8.3 Optimization	15
8.3.1 Relax and penalize: a new bilevel approach to mixed-binary hyperparameter optimization (TMLR'25) [11]	15
8.4 Other major works	15
8.4.1 Self-Play Q-Learners Can Provably Collude in the Iterated Prisoner's Dilemma (ICML'25) [12]	15

9	Bilateral contracts and grants with industry	16
9.1	Bilateral contracts with industry	16
9.1.1	Partnership Inria-IFPEN (2025-2028)	16
9.1.2	CIFRE Theses with Thalès (2024-2027)	16
9.1.3	I-Démo Région "GREENAI" (2024-2027)	16
9.2	Bilateral Grants with Industry	17
9.2.1	"Baby Cry" project - AXA Foundation (2024-2026)	17
10	Partnerships and cooperations	17
10.1	International initiatives	17
10.1.1	Inria associate team not involved in an IIL or an international program	17
10.2	International research visitors	18
10.2.1	Visits of international scientists	18
10.2.2	Visits to international teams	18
10.3	European initiatives	19
10.4	National initiatives	19
10.4.1	PEPR-IA PRODIGE-AI (2026-2030)	19
10.4.2	ANR MONALISA (2025-2029)	20
10.4.3	ANR LSD (2025-2028)	20
10.4.4	ANR MELISSA (2024-2029)	21
10.4.5	ANR DATES (2025-2029)	21
10.4.6	ANR ACOULAK (2025-2029)	22
10.4.7	ANR TAUDoS (2021-2026)	22
10.4.8	ANR SAFE (2022-2026)	23
10.4.9	ANR FAMOUS (2023-2027)	24
10.4.10	EUR SLEIGHT TREASURF (2024-2027)	24
10.4.11	EUR SLEIGHT PROXIMA (2025-2027)	25
11	Dissemination	25
11.1	Promoting scientific activities	25
11.1.1	Scientific events: organisation	25
11.1.2	Scientific events: selection	26
11.1.3	Journal	26
11.1.4	Invited talks	27
11.1.5	Scientific expertise	27
11.1.6	Research administration	27
11.2	Teaching - Supervision - Juries - Educational and pedagogical outreach	28
11.2.1	Teaching	28
11.2.2	Supervision	30
11.2.3	Juries	30
11.3	Popularization	31
11.3.1	Productions (articles, videos, podcasts, serious games, ...)	31
11.3.2	Participation in Live events	31
11.3.3	Others science outreach relevant activities	31
12	Scientific production	32
12.1	Major publications	32
12.2	Publications of the year	32

1 Team members, visitors, external collaborators

Research Scientists

- Quentin Bertrand [INRIA, Researcher]
- Benjamin Girault [INRIA, ISFP]

Faculty Members

- Marc Sebban [Team leader, UNIV Jean Monnet, Professor]
- Eduardo Brandao [UNIV Jean Monnet, Associate Professor]
- Farah Cherfaoui [UNIV Jean Monnet, Associate Professor]
- Rémi Emonet [UNIV Jean Monnet, Professor, HDR]
- Rémi Eyraud [UNIV Jean Monnet, Associate Professor Delegation, from Sep 2025, HDR]
- Rémi Eyraud [UNIV Jean Monnet, Associate Professor, until Aug 2025, HDR]
- Jordan Frecon Patracone [UNIV Jean Monnet, Associate Professor]
- Amaury Habrard [UNIV Jean Monnet, Professor, HDR]

Post-Doctoral Fellows

- Antoine Caradot [UNIV Jean Monnet]
- Chiheb Daaloul [UNIV Jean Monnet, Post-Doctoral Fellow, from Sep 2025]
- Volodimir Mitarchuk [UNIV Jean Monnet, Post-Doctoral Fellow, until Aug 2025]

PhD Students

- Fayad Ali Banna [UNIV Jean Monnet]
- Hind Atbir [UNIV Jean Monnet]
- Sayan Chaki [UNIV Jean Monnet]
- Ben Gao [UNIV Jean Monnet]
- Mickaël Gault [IFPEN, CIFRE, from Oct 2025]
- Thibault Girardin [UNIV Jean Monnet]
- Erick Gomez [UNIV Jean Monnet]
- Mael Jousset [UNIV Jean Monnet, from Oct 2025]
- Petros Kafkas [UNIV Jean Monnet, from Oct 2025]
- Dorian Llavata [UNIV Jean Monnet]
- Robin Mermillod-Blondin [HID Global]
- Abdel-Rahim Mezidi [UNIV Jean Monnet]
- Diego Pinto-Suarez [UNIV Jean Monnet, from Dec 2025]
- Joan Roux [UNIV Jean Monnet, from Oct 2025]

Interns and Apprentices

- Hasti Bargharariyan [UNIV Jean Monnet, Intern, from Mar 2025 until Aug 2025]
- Patrick Barry [UNIV Jean Monnet, Intern, from Apr 2025 until Aug 2025]
- Circée Chalayer [UNIV Jean Monnet, Intern, from Apr 2025 until Jul 2025]
- Alexin Choisnet [UNIV Jean Monnet, Intern, from May 2025 until Jul 2025]
- Eduard Andrei Duta Costache [INRIA, Intern, from Apr 2025 until Jul 2025]
- Emeric Gandon [UNIV Jean Monnet, Intern, from Sep 2025]
- Valentin Malquy [UNIV Jean Monnet, Intern, from May 2025 until Jul 2025]
- Mahima Sumathi [UNIV Jean Monnet, Intern, from Apr 2025 until Aug 2025]

Administrative Assistants

- Sylvie Boyer [Inria]
- Naima Chalais Traore [UNIV Jean Monnet]

2 Overall objectives

Our Inria MALICE team, whose members have a strong expertise in statistical learning, applied mathematics, statistics and optimization, develops algorithmic and theoretical research focused on integrating physical knowledge into machine learning (ML) models. Leveraging the skills present at the Hubert Curien lab in physics, MALICE aims to foster the development of new methodological contributions in Physics-informed Machine Learning (PiML) with a primary targeted application in Surface Engineering, making possible scientific breakthroughs in both Machine Learning and Physics. Our team focuses on several challenges, including (i) scarce observation data and incomplete background knowledge (typically in the form of Partial Differential Equations - PDEs), (ii) the need of deriving theoretical (generalization, approximation, optimization) guarantees on models learned from both data and physical knowledge and (iii) a strong necessity to transfer knowledge from one dynamical system to another. The advances carried out in machine learning allow to better understand the physics underlying the mechanisms of laser/radiation-matter interaction, enabling to address numerous societal challenges in the fields of space, nuclear, defense, energy or health.

3 Research program

MALICE is rooted at the interface of applied mathematics, statistical learning theory, optimization, physics and differentiable simulation. The following three scientific axes aim at addressing the aforementioned challenges from both theoretical (Axis 1) and algorithmic (Axis 2), as well as applied (Axis 3) perspectives.

Axis 1: Theoretical Frameworks when learning from data and background knowledge

Generalization guarantees typically aim at bounding the deviation of the true risk of an hypothesis from its empirical counterpart. These bounds, often referred to as PAC (Probably Approximately Correct) bounds, are usually derived by resorting to concentration inequalities (*e.g.* Chebyshev, Hoeffding, McDiarmid, etc.). Several theoretical frameworks have been introduced in the literature for establishing generalization bounds, including uniform convergence, uniform stability, algorithmic robustness or PAC-Bayesian theory, to cite a few. The state of the art bounds differ in the way (i) they incorporate some complexity measure (*e.g.* VC-dimension, Rademacher complexity, fat-shattering dimension, uniform stability constant, covering number, divergence, etc.) and (ii) take into account how the learning algorithm searches (or not) the parameter space. The ambitious goal of this Axis is to investigate how generalization bounds can be derived when the learning algorithm has access to both training data (simulation and/or observation examples) and background knowledge (typically in the form of PDEs). Addressing this task raises several challenges: How to define

complexity measures that capture dynamics' characteristics of the physical models? Can we choose the simulation data that makes bounds tighter? Can we derive such bounds in a transfer learning setting where two different but related dynamics are involved? As bilevel optimization seems to be very promising for addressing tasks involving data and knowledge, we also investigate how to derive statistical guarantees of nonsmooth bilevel problems. On the other hand, it is worth noticing that most of the physics-informed machine learning methods, like PINNs, minimize, as a term of the loss function, the residuals of the PDE. In this context, the team aims to study the approximation guarantees of such neural networks so as to ensure that minimizing the residuals leads to a small prediction error on the target vector field. Such guarantees are key in laser-matter interaction where the dynamics is governed by complex high-order non linear PDEs, like the Swift-Hohenberg or Kuramoto–Sivashinsky equations.

Note that the methodological contributions of this Axis aim to guide the design of the algorithms developed in Axes 2 and 3.

Tools and methods used in this axis: *complexity measures, PAC-Bayes theory, algorithmic robustness, uniform stability, concentration inequalities, bilevel optimization, functional analysis, Sobolev norms.*

Axis 2: Integration and extraction of knowledge in Physics-informed ML

Embedding physical laws in learning processes and discovering/augmenting knowledge from data are particularly difficult tasks in surface engineering which is governed by both low data and low knowledge regimes, *i.e.* where (i) the amount of observation data at our disposal is limited because of heavy experimental setups, (ii) based on tightly intertwined dynamics, the available physical theories (typically PDEs) describe partially the underlying phenomenon and finally, (iii) the actual continuous dynamics is not observable leading to unregistered training data pairs. Indeed, unlike other dynamical systems (e.g. in meteorology, temperature lake modeling, fluid mechanics, etc.), because light propagates too fast and the interaction with the matter lasts only a few femtoseconds (preventing any optical devices from taking images), we do not have easily access with the same initial conditions to the states of the system at time $t, t + \delta_t, \dots, t + n\delta_t$. Among the research avenues that are explored to address this strongly constrained scenario, we aim to design hybrid (data+knowledge) methods for benefiting from the best of the two worlds.

Another line of research consists in studying how bilevel optimization can integrate into two nested levels both the data and the surface engineering knowledge, and investigating (i) how sparse modeling might overcome the lack of data and (ii) how to augment/discover the underlying physical knowledge by seeing the orders of the partial derivatives as (real) hyperparameters. We also attempt to circumvent the problem of data scarcity by exploring ways to better estimate time derivatives using unrolled methods between two successive time steps. Finally, in this low data regime characterizing surface engineering, a special focus is placed on generative models which may turn out to be highly useful for data augmentation.

Tools and methods used in this axis: *sparse modeling, PDE learning, automatic differentiation, generative AI, bilevel optimization, neural operators, numerical schemes.*

Axis 3: Domain Generalization and Transfer Learning for Surface Engineering

Standard machine learning algorithms work well under the common assumption that the training and test data are drawn according to the same distribution. When the latter changes at test time, most statistical models must be reconstructed from newly collected data, which for some applications can be costly or impossible to obtain. Therefore, it has become necessary to develop approaches that reduce the need and the effort to obtain new labeled samples by exploiting pre-trained models and/or data that are available in related areas, and using these further across similar fields. This has given rise to the transfer learning and domain generalization frameworks which received a tremendous interest in the past years from the machine learning community. A key challenge in the MALICE project is to develop novel transfer learning methods for surface engineering that are able to adapt to a change of physical context (matter properties, governing PDEs, spatio-temporal conditions, initial/boundary conditions, etc.). Beyond foundation models in surface engineering that will be examined during the project, the objective of this axis is to define new differentiable measures of discrepancy (possibly at different scales) able to capture these shifts of background knowledge. We also study the links that can be established between diffusion models, flow matching, optimal transport and domain adaptation.

Tools and methods used in this axis: *knowledge divergences, generative AI, flow matching, diffusion, optimal transport, cross-knowledge modeling, entropy of dynamical systems, foundation models.*

4 Application domains

The scientific advances that are carried out in the team can be directly exploitable in numerous applications related to surface engineering and laser matter interaction, including automotive sector, health, biology, medicine, micro-surgery, environment, energy, security, space, to cite a few. This is made possible by texturing and thus giving a function to the surface of the matter. Mastering this physics allows to enhance the way one can give specific properties to a material, e.g. to obtain different optical effects of a glass according to the field of observation (reflection, transmission), to improve image visual rendering by laser-induced printing, to structure or texture tissues to control some biological behavior or to adapt surface matter to control the friction or adherence. For instance, the control of the nanopeak organization would open the door to advances in the protection of materials (living or not) against the attacks of bacteria or virus. On the other hand, nanocavities might be of great interest in the nanostructuring of car cylinders leading to CO₂ emission reduction. The scientific breakthroughs achieved in MALICE aim at strengthening the existing collaborations between the laboratory and major economic and public stakeholders in the fields of space, nuclear, defense, energy, automotive, health, including CNES, CERN, ORANO, CEA, DGA, EXAIL, HEF, RENAULT, BIOMERIEUX, HID GLOBAL, THALES, ST MICROELECTRONICS, CETIM, IFPEN, to cite some of them.

5 Social and environmental responsibility

One objective of the team is to show that it is possible to learn (theoretically) well from less data by leveraging physical knowledge. The assumption is that the latter can play the role of regularization leading to the prediction of plausible solutions while allowing a reduction of the number of examples required to train neural networks. This might have an impact on the carbon footprint by reducing the amount of costly collected data (involving matter irradiation and microscopy imaging) and therefore the computational resources needed for training the models. On the other hand, we also optimize surrogate neural solvers that allow to explore in a cheaper way the feature space and avoid running countless costly numerical simulations.

6 Highlights of the year

6.1 Papers

The following four papers, illustrating some of the scientific highlights of our team, were accepted at NeurIPS and ICML 2025 (see more details in Section 8).

- On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity [2] (selected as an oral - top 0.3% - at NeurIPS 2025).
- Conformal Online Learning of Deep Koopman Linear Embeddings (NeurIPS'25) [5]
- A Bregman Proximal Viewpoint on Neural Operators (ICML'2025) [6]
- Self-Play Q-Learners Can Provably Collude in the Iterated Prisoner's Dilemma (ICML'2025) [1]

A last paper (entitled "Photonic Learning in Ultrafast Laser-Induced Complexity"), submitted in 2025 and just accepted in January 2026, highlights a major milestone in our collaborations with physicists and has been accepted in the highly selective international journal Ultrafast Science.

These papers demonstrate the team's ability to publish at the highest level not only in machine learning but also in physics.

6.2 Inria Associate Teams

MALICE initiated the creation of two Inria associate teams in 2025 (see more details in Section 10): DYNAMO, led by Jordan Patracone in collaboration with IIT (Italy), and LSD, led by Quentin Bertrand in collaboration with MILA (Canada).

6.3 Nomination

Quentin Bertrand is an affiliated member with MILA (Canada) for three years since September 19, 2025.

6.4 Awards

- Amaury Habrard: ICML'25 outstanding area chair (top 2%).
- Quentin Bertrand: NeurIPS'25 top-reviewer.

7 Latest software developments, platforms, open data

7.1 Latest software developments

7.1.1 SwiftHohenbergPseudoSpectral

Keywords: Partial differential equation, Numerical solver, Self-organization

Functional Description: This software solves the dimensionless Swift-Hohenberg Equation with a cubic-quadratic nonlinear term. The Swift-Hohenberg Equation is a widely studied model for pattern formation, derived using symmetry arguments. As such, it serves as a maximally symmetric model for systems exhibiting spontaneous pattern formation, making it broadly applicable to various physical contexts.

The solver employs pseudospectral methods based on Fourier representations to ensure efficient and accurate computation. Importantly, the solver is physically constrained, using a Lyapunov functional to validate solutions, ensuring that the computed results remain consistent with the physical behavior expected from the system.

Designed with versatility and automation in mind, the solver is capable of generating a large number of solutions with minimal supervision. This feature makes it particularly well-suited for creating synthetic datasets to be used in Physics Guided Machine Learning. Specifically, it addresses problems related to the femtosecond laser-induced self-organization of matter, where data augmentation is crucial for improving the performance and generalization of machine learning models.

Publications: [ujm-04157829](#), [ujm-03823722](#), [tel-04515418v1](#)

Contact: Eduardo Brandao

7.1.2 GUAP

Name: Generalized Universal Adversarial Perturbations

Keyword: Machine learning

Functional Description: The software extends universal attacks by jointly learning a set of perturbations to choose from, aiming to maximize the success rate of attacks against deep neural network models.

URL: <https://github.com/JordanFrecon/guap>

Contact: Jordan Frecon Patracone

Partner: LITIS

7.1.3 GraSPy

Name: Graph Signal Processing for Python

Keywords: Graph, Signal processing, Machine learning

Functional Description: Implements classic components of such a toolbox, including (weighted) graph generalization and transformation and of their signals (a.k.a. vertex attributes), specific data visualization using Plotly, signal transformation with spectral methods (using IAGFTs), and graph learning.

URL: <https://gitlab.inria.fr/begiraul/joint-graph-learning>

Publication: [hal-04025968](#)

Contact: Benjamin Girault

7.1.4 Scikit-SpLearn

Name: Toolbox for the spectral learning of weighted automata

Keywords: Machine learning, Weighted automata

Scientific Description: The core idea of the spectral learning of weighted automata is to use a rank factorization of a complete sub-block of the Hankel matrix of a target series to induce a weighted automaton.

Functional Description: The toolkit contains an efficient implementation of weighted automata, a library to transform any dataset containing strings of different length into the scikit-learn format for data, 4 variants of the spectral learning algorithm.

URL: <https://remieyraud.github.io/scikit-splearn/>

Publication: [hal-01777169](#)

Contact: Rémi Eyraud

Partner: Laboratoire d'Informatique et des Systèmes (LIS) Université Aix-Marseille

7.1.5 SoundScapeExplorer

Keywords: Acoustics, Data analysis, Data visualization, Soundscape

Functional Description: SSE stands for SoundScapeExplorer and works with passive acoustic monitoring campaigns made of several microphones recording during hours, days or week.

It analyzes these data by extracting various features and indicators for each signal window (typically 1 second) then aggregating them by chunks (e.g. from 5 seconds to 1 or more minutes). Using a selection of dimensionality reduction techniques, it allows to visualize the whole campaign as a point cloud where each point corresponds to a temporal chunk on one microphone. The visualization is interactive, allowing coloring by criteria, filtering, listening to audio data, cluster analysis and beyond.

URL: <https://sound-scape-explorer.github.io/>

Contact: Rémi Emonet

Partner: Laboratoire ENES

8 New results

During 2025, the MALICE team contributed to the various areas of its scientific project, notably achieving new results in the fields of generative AI, physics-informed machine learning, and bi-level optimization. These results led to publications not only in major machine learning conferences (NeurIPS, ICML, ICLR, ECML) but also in international conferences and journals in Physics. Once again this year, the ambitious objective of publishing in both communities was achieved during 2025.

8.1 Generative AI - Flow Matching

In the context of surface engineering, where data is scarce, generative AI can be highly useful for better understanding the underlying physical dynamics, performing data augmentation (Axis 2) or addressing domain adaptation tasks (Axis 3). In collaboration with the OCKHAM Inria team, we have carried out in 2025 research in the field of flow matching leading to an oral presentation at NeurIPS'25 and a blogpost at ICLR'25.

8.1.1 On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity (NeurIPS'25) [2]

Participants: Quentin Bertrand, Rémi Emonet.

Collaboration with Anne Gagneux and Mathurin Massias (Inria OCKHAM)

Modern deep generative models can now produce high-quality synthetic samples that are often indistinguishable from real training data. A growing body of research aims to understand why recent methods—such as diffusion and flow matching techniques—generalize so effectively. Among the proposed explanations are the inductive biases of deep learning architectures and the stochastic nature of the conditional flow matching loss. In this work, we rule out the latter -the noisy nature of the loss -as a primary contributor to generalization in flow matching. First, we empirically show that in high-dimensional settings, the stochastic and closed-form versions of the flow matching loss yield nearly equivalent losses. Then, using state-of-the-art flow matching models on standard image datasets, we demonstrate that both variants achieve comparable statistical performance, with the surprising observation that using the closed-form can even improve performance.

Keywords: Generative AI, Flow Matching

8.1.2 A Visual Dive into Conditional Flow Matching (ICLR Blogposts'25) [4]

Participants: Quentin Bertrand, Rémi Emonet.

Collaboration with Ségolène Martin (Technische Universität Berlin), Anne Gagneux and Mathurin Massias (Inria OCKHAM)

Conditional flow matching (CFM) was introduced by three simultaneous papers at ICLR 2023, through different approaches (conditional matching, rectifying flows and stochastic interpolants). The main part of this post, Section 2, explains CFM by using both visual intuitions and insights on its probabilistic formulations. Section 1 introduces normalizing flows; it can be skipped by reader familiar with the topic, or that wants to cover them later. Section 3 opens on the links between CFM and other approaches, and ends with a 'CFM playground'.

Keywords: Generative AI, Flow Matching

8.2 Physics-informed Machine Learning (PIML)

The following works have contributed to both theoretical and algorithmic advances in Physics-informed Machine Learning for surface engineering (Axis 1 and 2 of our scientific objectives) as well as in learning nonlinear dynamical systems. These contributions have been published in 2025 in top-tier conferences in machine learning (i.e. NeurIPS, ICML, ECML) and journals in physics.

8.2.1 A Bregman Proximal Viewpoint on Neural Operators (ICML'25) [6]

Participants: Abdel-Rahim Mezidi, Jordan Patracone, Amaury Habrard, Rémi Emonet, Marc Sebban.

Collaboration with Saverio Salzo (Sapienza University of Rome, Italy)

We present several advances on neural operators by viewing the action of operator layers as the minimizers of Bregman regularized optimization problems over Banach function spaces. The proposed framework allows interpreting the activation operators as Bregman proximity operators from dual to primal space. This novel viewpoint is general enough to recover classical neural operators as well as a new variant, coined Bregman neural operators, which includes the inverse activation operator and features the same expressivity of standard neural operators. Numerical experiments support the added benefits of the Bregman variant of Fourier neural operators for training deeper and more accurate models.

Keywords: Neural Operators, Physics-informed ML

8.2.2 Provably Accurate Adaptive Sampling for Collocation Points in Physics-informed Neural Networks (ECML'25) [3]

Participants: Antoine Caradot, Rémi Emonet, Abdel-Rahim Mezidi, Amaury Habrard, Marc Sebban.

Despite considerable scientific advances in numerical simulation, efficiently solving PDEs remains a complex and often expensive problem. Physics-informed Neural Networks (PINN) have emerged as an efficient way to learn surrogate solvers by embedding the PDE in the loss function and minimizing its residuals using automatic differentiation at so-called collocation points. Originally uniformly sampled, the choice of the latter has been the subject of recent advances leading to adaptive sampling refinements for PINNs. In this work, leveraging a new quadrature method for approximating definite integrals, we introduce a provably accurate sampling method for collocation points based on the Hessian of the PDE residuals. Comparative experiments conducted on a set of 1D and 2D PDEs demonstrate the benefits of our method.

Keywords: PINNS, Collocation points, Physics-informed ML

8.2.3 Conformal Online Learning of Deep Koopman Linear Embeddings (NeurIPS'25) [5]

Participants: Ben Gao, Jordan Patracone.

Collaboration with Stéphane Chrétien (ERIC-Lyon 2) and Olivier Alata (LabHC)

We introduce in this work Conformal Online Learning of Koopman embeddings (COLoKe), a novel framework for adaptively updating Koopman-invariant representations of nonlinear dynamical systems from streaming data. Our modeling approach combines deep feature learning with multi-step prediction consistency in the lifted space, where the dynamics evolve linearly. To prevent overfitting, COLoKe employs a

conformal-style mechanism that shifts the focus from evaluating the conformity of new states to assessing the consistency of the current Koopman model. Updates are triggered only when the current model's prediction error exceeds a calibrated threshold, allowing selective refinement of the Koopman operator and embedding. Empirical results on benchmark dynamical systems demonstrate the effectiveness of COLoKe in maintaining long-term predictive accuracy while significantly reducing unnecessary updates.

Keywords: Nonlinear dynamical systems, Streaming data.

8.2.4 Using Signatures and Koopman Operator to Learn Nonlinear Dynamics (GSI'25)

Participants: Ben Gao, Jordan Patracone.

Collaboration with Stéphane Chrétien (ERIC-Lyon 2) and Olivier Alata (LabHC)

We propose a novel framework for predicting the evolution of dynamical systems by learning the Koopman operator in the space of linear functionals on the Signature transform of trajectory data. The Signature, a central object in rough path theory, provides a universal and compact representation of paths through iterated integrals, enabling linear models to approximate a wide class of non-linear functionals. By restricting observables to lie in the span of truncated Signatures, we construct a finite-dimensional approximation of the Koopman operator, which we estimate directly from data using regularized linear regression. This approach merges the expressiveness of operator-theoretic methods with the structural richness of Signature features.

Keywords: Nonlinear dynamical systems, Signature transform.

8.2.5 Contextual Hypernetwork for Adaptive Prediction of Laser-Induced Colors on Quasi-Random Plasmonic Metasurfaces (ECML'25) [25]

Participants: Thibault Girardin, Amaury Habrard.

Collaboration with Nathalie Destouches (LabHC)

Laser processing is a rapid, versatile, and low-cost technology to print images on large surfaces. When applied to very thin films embedded with disordered metallic nanoparticles, known as quasi-random plasmonic metasurfaces, it generates colors that vary with the observation mode, making it valuable for visual security applications. Predicting these colors in different modes from the knowledge of laser processing parameters and the initial state of the metasurface can accelerate the industrialization process. However, there is no general physical model able to make this prediction accurately in various modes. In order to address this issue, this paper proposes a data-driven approach for learning deep models on experimental data able to predict the colors observed in different environments for a large range of laser processing parameters. We leverage a framework that learns jointly a shared latent space for multiple environments together with a contextual representation specific to each. This contextual representation is generated by a hypernetwork conditioned on an interpretable context vector. This context vector can be learned from few data allowing fast adaptation to new environments. This approach demonstrates that a single model can learn to predict a large range of colors across different environments. Its effectiveness is demonstrated through its ability to rapidly adapt to new scenarios with minimal data and to serve as an improved weight initializer for finetuning when larger datasets are available. Source code and datasets are available on Gitlab.

Keywords: Laser-matter interaction, Domain Adaptation, Hypernetwork

8.2.6 Physics-Informed Machine Learning for Modeling CO₂ Capture from Scarce Data (ICTAI'25) [18]

Participants: Mickael Gault, Rémi Emonet, Marc Sebban.

Collaboration with P. Bachaud, B. Celse (IFPEN, France)

Accurate modeling of complex industrial processes often relies on costly mechanistic simulations grounded in physical principles. In this paper, we investigate the subject of CO₂ capture, a major environmental challenge, through the absorption column of an amine-based post-combustion process. The modeling of such unit at industrial scale faces two difficulties: (i) theoretical models, efficient at laboratory scale, might fail to fully reflect the complexity of the numerous intertwined phenomena occurring in the absorber, (ii) the cost and uncertainty of industrial observation data make purely data-driven approaches unfeasible.

To tackle both this low data regime and inaccurate physical models, we envision this CO₂ capture problem through the lens of Physics-informed Machine Learning (PiML). We present a hybrid (data+knowledge) model where the scarce observation data complement the physical model, while the latter ensures that the predictions remain physically consistent. Beyond the standard use of simulation data for learning and the embedding of physical laws as regularization, the originality of our PiML algorithm compared to other methods in the literature lies in a physical prior assumption about the network architecture and its countercurrent flow learning process inspired by the column's operation. Our experimental results showcase a significant improvement in accuracy and highlight the potential of our augmented model for generalizing across domains, especially when data is scarce.

Keywords: Physics-informed ML

8.2.7 Image quality metrics for restricted gamut images produced by laser-induced printing on plasmonic films (Journal of the European Optical Society 2025) [9]

Participants: Rémi Emonet.

Collaboration with Robin Mermillod-Blondin, Nicolas Dalloz, Alain Tremeau, Aldi Wista Fadhillah, and Nathalie Destouches (LabHC)

Laser-induced printing is a low-cost, high-speed, non-contact method of marking large, high-resolution images. Implemented on thin films containing metallic nanoparticles, the technique allows for the printing of color images with visual effects. However, these images typically have a limited color gamut compared to inkjet printing. This limitation is due to the inability to achieve high levels of saturation for all colors and to cover the sRGB hue range. While common quality metrics focus primarily on aspects such as resolution or blur, they rarely address the color aspect. This study proposes a methodology to provide image quality metrics adapted to color gamuts with unusual shapes and volumes. It aims to rank them in terms of image quality performance for any given image. In particular, this work focuses on gamuts measured in transmission and reflection that are not necessarily centered on the CIE a^*b^* plane and may exhibit low contrast. Psychophysical studies have been conducted to evaluate the quality of images simulated with different color gamuts. The same images were evaluated using different metrics, and an analysis based on the ANOVA model was used to determine a set of metrics that explain observers' preferences.

Keywords: Laser-matter interaction, Machine Learning

Note that other works in physics-informed machine learning have been presented at physics conferences, such as [21, 26].

8.3 Optimization

In MALICE, we study bilevel optimization as a promising framework for addressing physics-based machine learning tasks, where the knowledge and the data would be involved into two different nested levels. We are also interested in leveraging bilevel optimization for hyper-parameter selection (Axis 1 and 2). We published this year in the international journal TMLR a new bilevel approach to mixed-binary hyperparameter optimization. This work was carried out in particular with Saverio Salzo, a partner of the associate Inria team Dynamo led by Jordan Patracone.

8.3.1 Relax and penalize: a new bilevel approach to mixed-binary hyperparameter optimization (TMLR'25) [11]

Participants: Jordan Patracone.

Collaboration with Sara Venturini, Marianna de Santis, Francesco Rinaldi, Saverio Salzo (Italy), Martin Schmidt (Germany)

In recent years, bilevel approaches have become very popular to efficiently estimate high-dimensional hyperparameters of machine learning models. However, to date, binary parameters are handled by continuous relaxation and rounding strategies, which could lead to inconsistent solutions. In this context, we tackle the challenging optimization of mixed-binary hyperparameters by resorting to an equivalent continuous bilevel reformulation based on an appropriate penalty term. We propose an algorithmic framework that, under suitable assumptions, is guaranteed to provide mixed-binary solutions. Moreover, the generality of the method allows to safely use existing continuous bilevel solvers within the proposed framework. We evaluate the performance of our approach for two specific machine learning problems, i.e., the estimation of the group-sparsity structure in regression problems and the data distillation problem. The reported results show that our method is competitive with state-of-the-art approaches based on relaxation and rounding.

Keywords: Bilevel Optimization

8.4 Other major works

The following work was carried out in particular with Gauthier Gaudel from MILA (Canada), a partner of the associate Inria team LSD led by Quentin Bertrand.

8.4.1 Self-Play Q-Learners Can Provably Collude in the Iterated Prisoner's Dilemma (ICML'25) [12]

Participants: Quentin Bertrand.

Collaboration with Emilio Calvano (LUISS Business School, Università LUISS Guido Carli, Rome, Italy), Juan Agustin Duque and Gauthier Gidel from MILA (Canada)

A growing body of computational studies shows that simple machine learning agents converge to cooperative behaviors in social dilemmas, such as collusive price-setting in oligopoly markets, raising questions about what drives this outcome. In this work, we provide theoretical foundations for this phenomenon in the context of self-play multi-agent Q-learners in the iterated prisoner's dilemma. We characterize broad conditions under which such agents provably learn the cooperative Pavlov (win-stay, lose-shift) policy rather than the Pareto-dominated "always defect" policy. We validate our theoretical results through additional experiments, demonstrating their robustness across a broader class of deep learning algorithms.

Keywords: Q-Learning, Cooperation, Prisoner dilemma

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

9.1.1 Partnership Inria-IFPEN (2025-2028)

Participants: Rémi Emonet, Marc Sebban.

Partner: IFPEN Solaize - Conception, Modélisation, Procédés.

This newly signed 2025-2028 project is part of the national strategic partnership between Inria and IFPEN in support of the energy transition, which led to the creation of a joint laboratory between the two partners in 2020. It deals with Physics-informed Machine Learning, especially with hybrid deep-learning models that combines physical laws with data-driven approaches. Their promise is to be more interpretable, more robust when facing outliers, and to lead to a more plausible physical behavior than their pure-machine-learning counterparts. The application case is post-combustion carbon capture, a process of major importance to tackle hard-to-abate greenhouse gas emissions of some key industrial sectors. The absorption column of this type of unit is modeled using a rigorous description of the physical mechanisms responsible for the transfer of acid gas to the liquid solvent. Nonetheless, this model fails to predict some experimental tendencies observed on industrial units, and these discrepancies cannot be explained by a specific phenomenon. In this context, the objective of this joint project (including a PhD thesis co-supervised by Rémi Emonet and Marc Sebban) is to design new hybrid PiML approaches, which incorporate laws from chemistry, to improve the precision of the absorption column description.

9.1.2 CIFRE Theses with Thalès (2024-2027)

Participants: Rémi Eyraud.

Partners: SESAM (LabHC-UJM), Thalès-DIS

During Fall 2024, a collaboration with the DIS (Digital Identity & Security) lab of the Thalès firm started via a CIFRE PhD (Wissal Ghamour). Rémi Eyraud took the academic direction of the PhD together with members of the SESAME team of the Hubert Curien Laboratory. The subject aims at leveraging the tools of Machine Learning for Side Channel Attacks of embedded microchips. Information from physics can be added to the process, since the leak mainly relies on an electro-magnetic sensor. This followed the Ph.D of Dorian Llavata that started in 2022 on the same subject in collaboration with the CESTI Leti of the CEA.

9.1.3 I-Démo Région "GREENAI" (2024-2027)

Participants: Jordan Patracone.

Partners: Dracula Technologies, ASYGN.

The project I-Démo Région "GREENAI" involves three key actors working collaboratively towards sustainable Internet of Things (IoT) solutions. Dracula Technologies specializes in developing organic photovoltaic cells that harness ambient light to power IoT devices, aligning with environmental sustainability goals. ASYGN contributes by designing ultra-low-power hardware accelerators to enable advanced AI processing directly on IoT devices. The Laboratoire Hubert Curien (Jordan Patracone) focuses on energy-efficient artificial intelligence for computer vision, optimizing neural network architectures for low-resource hardware through two scientific theses (among which, that of Ben Gao). This work is sponsored by a public grant overseen by the Auvergne-Rhône-Alpes region, Grenoble Alpes Métropole, and BPIFrance.

9.2 Bilateral Grants with Industry

9.2.1 "Baby Cry" project - AXA Foundation (2024-2026)

Participants: Rémi Emonet.

Partners: ENES (UJM), SAINBIOSE-MoVE, CHU Saint-Étienne.

Lead by the ENES, the AXA Baby Cry 1000 project aims at recording 1000 babies, each for one day after they are born to help early diagnosis of cognitive development issues and compare with the development of premature babies. The project is a joint effort by three labs (ENES, Laboratoire Hubert Curien (Rémi Emonet), and SAINBIOSE-MoVE, CHU Saint-Étienne) and is funded by the AXA Foundation for a total of 1M€ and a duration of 3 years.

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Inria associate team not involved in an IIL or an international program

MALICE initiated the creation of two Inria associate teams in 2025: **DYNAMO**, led by Jordan Patracone in collaboration with *IIT (Italy)*, and **LSD**, headed by Quentin Bertrand in collaboration with *MILA (Canada)*.

DYNAMO: DYNAmical systems, Analysis, and Machine learning for self-Organization of matter

Participants: Jordan Patracone (coordinator) , Eduardo Brandao, Amaury Habrard, Marc Sebban.

Partners: Massimiliano Pontil (Istituto Italiano di Tecnologia, Genoa, Italy); Saverio Salzo (Sapienza Università di Roma, Rome, Italy)

Type: Inria Associate Team

Duration: 3 years (2025-2028)

The objective of DYNAMO is to **advance the collaboration between the two partners with an ambitious project focused on machine learning & self-organization systems**. This project **leverages the unique strengths of both teams**, enabling us to tackle the associated complex challenges more effectively. Both teams share expertise in statistical learning theory, optimal transport, transfer learning, approximation theory, and bi-level optimization, which positions us well for success. The MALICE Inria team contributes its deep knowledge of physics modeling and physics-informed machine learning, while the Main Team offers expertise in stochastic processes, online learning, and kernel methods. Additionally, this initiative will allow us to **engage new collaborators from both the Italian and French sides**, further enriching our joint research efforts.

LSD: Leveraging Synthetic Data from Generative Models

Participants: Quentin Bertrand (coordinator) , Rémi Emonet, Marc Sebban.

Partners: Gauthier GIDEL, MILA (Canada)

Type: Inria Associate Team

Duration: 3 years (2025-2028)

In this collaboration, our goal is to investigate both the risks (model collapse) and the opportunities (unlimited data access) to provide a clearer understanding of when and how retraining on synthetic data can be a safe and effective approach. Specifically, the goal is to analyze how, and under what conditions, synthetically generated data can be beneficial, potentially acting as a form of regularization for models trained on the same dataset that the generative model was originally trained on (Karras et al., 2024). Interestingly, negative guidance from models trained on synthetic data can even improve the generative model's own performance (Alemohammad et al., 2024a). The second goal is to explore the utility of generative models for data augmentation in low-data scenarios, such as in clinical trials (Nasimzada et al., 2024), protein design (Huguet et al., 2024), or physical simulations, where the model is pre-trained on a larger, different dataset. The ultimate goal is to study the dynamics of generative models in multi-agent settings, particularly examining the behavior of Large Language Models when they interact in shared environments.

10.2 International research visitors

10.2.1 Visits of international scientists

Jorge Azorin

Status: Researcher

Institution of origin: University of Alicante

Country: Spain

Dates: July 2025 (one month)

Mobility program: Research stay funded by "Invited Professors" UJM grant

10.2.2 Visits to international teams

Quentin Bertrand

Visited institution: Mila and Université de Montréal

Country: Canada

Dates: July-August 2025 (2 months)

Mobility program/type of mobility: LSD Inria associated team

Quentin Bertrand

Visited institution: Mila and Université de Montréal

Country: Canada

Dates: December 2025 (1 month)

Mobility program/type of mobility: LSD Inria associated team

Marc Sebban

Visited institution: University of Alicante

Country: Spain

Dates: June 2025 (10 days)

Mobility program/type of mobility: Research stay funded the [T4EU european alliance](#).

Amaury Habrard

Visited Institution: Isaac Newton Institute of Mathematical Science, Cambridge

Country: United Kingdom

Dates: 3 weeks in May-July 2025

Mobility program/type of mobility: Representing, calibrating & leveraging prediction uncertainty from statistics to machine learning supported by EPSRC

10.3 European initiatives

ML4Health - Transform4Europe (2024-2026)

Participants: Amaury Habrard, Marc Sebban (coordinator) .

Partners: University of Alicante, Dpto.Tecnología Informática y Computación (DTIC)

Type: Seed Funding **T4EU** (co-funded by the European Union)

Duration: 2 years (2024-2026)

The objective of ML4Health is two-fold: (i) from a training perspective, develop a double master degree in AI between the two T4EU partners, built from two existing complementary programmes (Machine Learning & Data Mining master track at UJM and Máster Universitario en Inteligencia Artificial at UA); (ii) from a scientific standpoint, benefit from the expertise in statistical learning and physics-informed Machine Learning (UJM) and in 3D perception and intelligent systems (DTIC) to focus on the morphological evolution of bodies of obese patients and the early detection of neurodegenerative diseases. In particular, the two partners study (data+theory) hybrid machine learning models combining both 3D+t images and biological/morphometric knowledge.

10.4 National initiatives

We describe below the main national projects we are involved in.

10.4.1 PEPR-IA PRODIGE-AI (2026-2030)

PRODIGE-AI: PROBability, ranDom matrIx theory, Geometry and gEneralization for generative-AI

Participants: Amaury Habrard (coordinator) , Quentin Bertrand, Farah Cherfaoui, Rémi Emonet, Rémi Eyraud, Benjamin Girault, Jordan Patracone, Marc Sebban.

Partners: LIS & IMT (Marseille), LJAD (Nice), LITIS (Rouen), LPT & IMT (Toulouse) & E. Morvant (LabHC-UJM)

Type: PEPR-IA project

The rapid and impressive development of generative AI opens up transformative projects for several key sectors such as healthcare, education and industry. This evolution comes with major challenges in order to ensure safe, ethical, and effective development of these technologies. In particular, ethical issues such as plagiarism and the amplification of existing biases are currently paramount. Guarantees regarding the functioning of these generative models and the understanding of their underlying mechanisms remain preliminary. The PRODIGE-AI project aims to address three crucial issues for the development of safer, more efficient, and more transparent generative AI: the generalization capabilities of generative AI models, the efficiency and explainability of these AI models, and the development of geometric generative AI - in particular for graphs. These three areas represent topics where the state of the art requires new foundational frameworks to enable

a controlled development. Our shared vision is that the promising potential of generative AI cannot be realized without a deep understanding of the fundamentals of generative deep neural networks and the mathematics underlying their learning mechanisms. Towards that endeavor, our consortium brings together both specialists in the foundations of machine learning and mathematicians experts in probability theory, graph theory and geometry. In mathematics, the project relies on expertise in stochastic processes, filtering, high dimensional statistics, random matrices, random graphs, random tensors, free probability, C^* -algebras, graph theory, classical and quantum information theory, and optimal transport. Another important objective of the project is to structure a group of researchers at the frontier between the foundations of machine learning and probability theory to foster the development of innovative and relevant results for generative AI.

10.4.2 ANR MONALISA (2025-2029)

MONALISA: Monotone Operators and Neural Architectures - Leveraging Interactions for physically Structured Approximations

Participants: Jordan Patracone (coordinator) .

Partners: CSML / IIT (Italy), DIAG / Univ. Sapienza di Roma (Italy), Univ. Novi Sad (Serbia), ERIC / Univ. Lyon 2

Type: ANR JCJC

Neural operators have recently shown great success in approximating complex, high-dimensional dynamical systems, such as those governed by partial differential equations (PDEs). Their advantage over traditional neural networks lies in their ability to model mappings between function spaces, capturing continuous dynamics across scales with discretization-invariant performance. Physics-informed neural operators further embed physical constraints, enforcing laws like PDEs. Trained with both data-driven and physics-based regularization, they offer faster predictions than traditional PDE solvers. However, the reasons behind their effectiveness remain unclear. A critical aspect of this ongoing exploration is determining how the complexity of these operators is controlled—whether through explicit regularization techniques or implicitly via their architectural design. In MONALISA, we aim to adopt an ambitious approach and borrow tools from the monotone operator theory, an area of nonlinear analysis recently applied to the study of nonlinear mappings modeled by neural architectures, to further understand and control the complexity of physics-informed neural operators. To implicitly control their complexity, the first research axis will leverage such connection to design neural operators whose architecture itself is guided by physical knowledge such as PDE. A starting point will be to formulate the action of an operator layer as the solution of a regularized optimization problem over function spaces. Then, we will study the impact of explicit regularization both at the layer-level as mentioned above and at the training loss level. Moreover, theoretical guarantees will be devised by hinging on monotone operator theory and on the theory of reproducing kernel Banach spaces. Finally, we will focus on the challenging application of laser-matter interaction, where structuring the neural operator with all available knowledge is essential to overcome the scarcity of data.

10.4.3 ANR LSD (2025-2028)

LSD: Leveraging Synthetic Data From Generative Models

Participants: Quentin Bertrand (coordinator) , Rémi Emonet, Amaury Habrard, Marc Sebban.

Partners: UdeM Mila (Canada), Inria OCKHAM (France)

Type: ANR PRCI

Generative models are machine learning models that learn and replicate the underlying structure of the data. The quality of the data they produce has reached a level that surpasses the human ability to differentiate

between real and synthetic data. This opens up the possibility of virtually unlimited access to realistic synthetic data, which could be leveraged for data augmentation, especially in situations where real-world data is scarce. In fields such as physics, clinical applications, and protein design, synthetic data can enrich datasets and enhance model generalization.

With the rise of indistinguishable synthetic content being generated and shared online, deployed systems now face the unprecedented challenge of managing synthetic data alongside authentic data. A growing concern in generative AI is "self-consuming" models, which are retrained on their own (previously generated) data. Over time, this recursive process can yield overfitting artifacts, accumulation of biases, or inaccuracies, ultimately causing critical model degradations (a.k.a., model collapse).

Leveraging the complementary expertise of the involved partners in generative modelling, optimization, and open-source software, the project aims to systematically investigate the risks and benefits of interactions between learning algorithms and synthetic data. One key objective is to assess when and to what extent generative models can improve performance on downstream tasks. Another goal is to quantify the rate at which "self-consuming" models collapse and to develop strategies to mitigate it. More broadly, the project will explore how generative models behave and interact when deployed in shared environments with multiple models or agents. This includes understanding how they influence each other's outputs, how they potentially cooperate or compete, and the impact of these interactions on overall system performance.

10.4.4 ANR MELISSA (2024-2029)

MELISSA: METHodological contributions in statistical Learning InSpired by SurfAce engineering

Participants: Eduardo Brandao, Rémi Emonet, Benjamin Girault, Amaury Habrard, Jordan Patracone, Marc Sebban (coordinator) .

Partners: MAGNET (Inria Lille), MLIA (ISIR, PSL) and F. Garrelie and JP Colombier (LabHC-UJM)

Type: ANR PRC

Website: [MELISSA](#)

The underlying dynamics of many physical problems are governed by parameterized partial differential equations (PDEs). Despite important scientific advances in numerical simulation, solving efficiently PDEs remains complex and often prohibitively expensive. Physics-informed Machine Learning (PiML) has recently emerged as a promising way to learn efficient surrogate solvers, and augment the physical laws by leveraging knowledge extracted from data. Despite indisputable advances, several open problems remain to be addressed in PIML: (i) Deriving generalization guarantees; (ii) Learning with a limited amount of data; (iii) Augmenting partially known physical laws; (v) Modeling uncertainty; (vi) Building foundation models for physics. MELISSA will deal with these problems from both theoretical and algorithmic perspectives. The objective is to design the next generation of provably accurate PIML algorithms in the challenging context of laser-matter interaction where data is scarce and the available physical laws only partially explain the observed dynamics.

10.4.5 ANR DATES (2025-2029)

DATES: Domain Adaptation on Time Series

Participants: Rémi Emonet (Principal coordinator) , Farah Cherfaoui, Quentin Bertrand, Marc Sebban.

Partners: IRISA (Rennes), ERICSSON (Firm)

Type: ANR PRCE

Machine learning models generally assume that training and test data are identically distributed. However, this assumption often breaks down in real-world applications due to shifting data distributions over time and across contexts. Domain Adaptation (DA) tackles this challenge by allowing models to leverage labeled data

from one domain to perform tasks in a different, often unlabeled, target domain. Time Series (TS) consists of data points recorded over time and is increasingly prevalent with autonomous sensors and online activities. In the context of DA, TS presents several challenges, including class imbalance, temporal drift, correlations and dependencies, and hierarchical structures that manifest as multi-scale periodic patterns. The DATeS project addresses three fundamental challenges of DA for TS. The first challenge is efficiently achieving both temporal and domain alignment, which we will tackle by proposing new formulations based on generalized optimal transport and conditional flow matching. The second challenge involves adapting to missing data and source-free domain adaptation, where only a learned model is available; we will tackle this in the lens of optimal transport approaches. The third challenge relates to unsupervised DA (with few or no target labels) and involves the general problem of selecting the best model, source domain(s), and hyperparameters without access to a validation set. We will address this challenge by learning metrics tailored for TS in DA and through meta-learning, enabling effective adaptation from multiple sources. This project aims to generate significant new knowledge in the field of DA for TS, addressing critical challenges that are currently under-explored. While DATeS is motivated by the challenges posed by time series data, it will also inspire innovative developments that have broader applicability beyond the TS domain.

10.4.6 ANR ACOULAK (2025-2029)

ACOULAK: Underwater acoustic pollution in peri-alpine lakes: Characterization and ecological impacts

Participants: Rémi Emonet, Marc Sebban.

Partners: ENES (UJM, Coordination), CARRTEL, INRA

Type: ANR PRC

Lakes are key resources for territories, providing many ecosystem services like water supply, aquatic biodiversity and human wellbeing. With the rise of recreational activities, lake managers worry about the increase in human presence on lakes with ever more powerful boats. One of the symptoms of water overcrowding is acoustic pollution. Boat sounds propagate in the air and generate conflicts between the populations of users, but also underwater where they are likely to disturb the aquatic fauna. Concern for freshwater acoustic pollution is very recent and ecosystem managers are seeking for assessment tools and fundamental knowledge on the ecological impacts to develop sustainable management strategies that promote human activities while preserving ecosystem integrity. We propose a timely and innovative project on underwater acoustic pollution using the peri-alpine lakes as model systems. Our first objective is to characterize underwater acoustic pollution in space and time using passive acoustic monitoring as new metrological tool and machine learning to process the data. We hypothesize that acoustic pollution varies in time depending on human activities and climatic variables, and in space according to local practices and regulations. Our second objective is to investigate impacts on dreissenid mussels and Eurasian perch larvae using playback experiments in the lab and the field. We hypothesize that increased sound level induces physiological and behavioral symptoms of a stress response. The consortium brings together experts from the fields of ecology, animal behavior, limnology and computer sciences. We will provide fundamental knowledge with international publications in the fields of environmental pollution, animal ecology, and machine learning, but also a unique database on freshwater underwater soundscapes, decision supports for lake managers, and materials for public awareness events.

10.4.7 ANR TAUDoS (2021-2026)

TAUDoS: Theory and Algorithms for the Understanding of Deep learning On Sequential data

Participants: Rémi Emonet, Rémi Eyraud (Principal coordinator), Amaury Habrard, Marc Sebban.

Partners: LIS (Aix-Marseille University), EURA NOVA (Firm), MILA (Canadian State)

Type: ANR PRCE

Website: [TAUDoS](#)

The ambition of this project is to provide a better understanding of the mechanisms that allow the amazing recent achievements of Machine Learning, and in particular of Deep Learning. This is achieved by providing elements that allow a better scientific comprehension of the models, strengthening our experimental results by theoretical guarantees, incorporating components dedicated to interpretability within the models, and allowing trustful quantitative comparison between learned models.

The originality and the specificities of TAUDoS are due to three major characteristics:

- The focus on models for sequential data, such as Recurrent Neural Networks (RNN), while most works concentrate on feed-forward networks
- The will to analyze these models in the light of formal language theory
- The goal to target both rigorous theoretical analyses and empirical evidence related to interpretability.

10.4.8 ANR SAFE (2022-2026)

SAFE: Controlling networks with safety bounded and interpretable machine learning

Participants: Amaury Habrard.

Partners: XLIM / Univ. Poitiers, IRISA / Univ. Rennes 1, Huawei, QOS DESIGN and B. Jeudy & K. Singh (LabHC-UJM)

Type: ANR PRCE

When applied to communication networks, traditional approaches for control and decision-making require a comprehensive knowledge of system and user behaviours, which is unrealistic in practice when there is an increase in scale and complexity. Data-driven AI approaches do not have this drawback, but offer no safety bounds and are difficult to interpret. The SAFE project aims to design an innovative approach by combining the best of both worlds. In this new approach, intelligence is distributed in the network between a global AI (at the central level) and local AIs (at the edge level) collaborating with each other by integrating traditional models with graph neural networks and reinforcement learning. The approach, developed for partially or completely observable/controllable environments, will natively integrate safety bounds, interpretability and provide self-adaptive systems for routing, traffic engineering and scheduling. SAFE has following scientific objectives with an open source strategy:

1. Hierarchical architecture: Assuming modern network architectures, we will design a ML architecture based on global AI (running at central controller level) and local AI (running at edge device level) for decision-making in partially as well as fully observable and controllable environments. Global AI will be able to control, configure and install policies on local AI.
2. Algorithms for partially observable environments: We will design new safety bounded and interpretable algorithms for self-adaptive traffic engineering, automatic scheduling algorithms for partially observable and controllable environments. These methods find use cases in SD-WAN (Software-Defined Wide Area Networks), where edge devices present at customer premises need to collaboratively operate in overlay on top of partially observable core networks.
3. Algorithms for fully observable environments: We will investigate the application of the global and local AI architecture for fully observable and controllable environments. Specifically, we will design new safety bounded and interpretable algorithms for software-defined routing and traffic engineering, which find use cases in data centers as well as private WANs connecting multiple sites.

10.4.9 ANR FAMOUS (2023-2027)

Famous: Fair Multi-modal Learning

Participants: Rémi Eyraud, Amaury Habrard.

Partners: LIS Aix-Marseille, LITIS Rouen, INT Marseille, Euranova and A. Gouru, C. Largeron & E. Morvant (LabHC-UJM)

The aim of this project is to explore the first avenues of research into the contribution of multimodality in datasets to meet the requirements of fair learning. Fairness refers here to the biases (in the data and/or induced), while being interested in the interpretability of the models to help their certification. Each modality has its own statistical and topological characteristics, which requires upstream research on the adjustment of distributions when biased, adapted metrics, etc. Moreover, each one being itself a bias of observation of the data, this will be taken into account to establish a joint distribution (trans-modal) unbiased on all these modalities. With theoretical research in cross-modal statistical learning, we will study methods for reducing some types of identified biases (non iid, imbalances, sensitive variables) in the case of multimodal data. Two levels of treatment are privileged: (1) cross-modal pre-processing of biases in the data, by learning metrics, neural representations, and optimization constraints on kernel pre-images; (2) cross-modal algorithms for eliminating biases in model learning: cross-modal optimization algorithms, as well as optimal transfer and transport approaches between modalities to debias the concerned ones, based on the theoretical results previously obtained. Parsimony will be considered for scaling and explainability. Transversally, our work will be based on problems arising from real data sets in biology and health, multi-modal and presenting various types of bias, and on toy data sets to be generated. They have modalities where the data are structured in graphs: all our fundamental works will be declined to take into account this specificity impacting the treatment of the considered biases.

10.4.10 EUR SLEIGHT TREASURF (2024-2027)

TREASURF: TRansfer IEarning for Frugal and Accurate modeling of SURface Functionalization prediction –application to multicomponent alloys

Participants: Amaury Habrard (coordinator), Rémi Emonet, Marc Sebban.

Partners: F. Garrelie and JP Colombier (LabHC-UJM)

Type: EUR Manutech SLEIGHT

TREASURF is an interdisciplinary project focusing on the development of novel machine learning approaches for the prediction of surface functionalization of different families of metals and metal alloys by (femto)laser irradiation. The ability to predict the micro- or nanopatterns induced by laser functionalization is a crucial challenge for an optimal use of surface properties. In this context, machine learning methods have been subject of a growing interest recently but they have to cope with limited amounts of experimental data due to the very high acquisition costs. In the TREASURF project, we propose to address this problem by developing methods able to transfer the knowledge of a prediction model learned from a given metal or alloy to another, different but sharing certain properties. Re-training a new model is not a plausible hypothesis, mainly because of the difficulties involved in acquiring large quantities of data (laser irradiation + nanoscale imaging). This project is therefore situated in a difficult context of "frugal" learning. Our aim is to focus primarily on topographic predictions for two or more different alloy families. The project also envisages taking into account variability due to chemical changes to guide the transfer process. The advances made in this project will enable us to better characterize the impact of laser-matter interaction with the perspective of designing new surface functionalizations on various novel metal alloys, opening the door to new application prospects in numerous societal challenges related to health, energy, space, nuclear or defense.

10.4.11 EUR SLEIGHT PROXIMA (2025-2027)

PROXIMA: Unfolded proximal neural operators for inverse problems in imaging

Participants: Jordan Patracone (co-coordinator) .

Partners: L. Denis (LabHC-UJM)

Type: EUR Manutech SLEIGHT

This project will develop a novel framework for solving inverse problems in imaging, focusing on the problem of joint reconstruction of a sequence of high-quality images from multiple degraded or incomplete observations. Classical approaches often rely on discretized formulations that are sensitive to resolution changes, leading to poor generalization across different scales. By operating directly in function spaces, our approach will ensure discretization invariance, allowing models trained on one resolution to generalize seamlessly to others. In this context, we will define and explore the extension of unfolded proximal networks – which design the deep network architecture itself by leveraging the variational formulation associated with the inverse problem – to function spaces, leading to unfolded proximal neural operators. This will allow for an adaptive and principled way to encode physical constraints and optimization-based priors directly into the learning process. Such an approach is particularly well-suited for ill-posed settings and applications where the sample undergoes unknown deformations, where classical data-driven methods often fail. A key challenge will be to extend generalization guarantees of unfolded networks beyond discrete settings, requiring new theoretical developments at the intersection of operator learning, proximal optimization, inverse problems, and statistical learning theory. We will also design self-supervised training strategies so that our methods can be applied to experimental data without requiring ground-truth images during the learning step (i.e., the clean and noise-free images corresponding to each actual observation). This framework will offer a scalable and interpretable solution for high-dimensional inverse problems in imaging, with potential applications in surface characterization or biological sample reconstruction.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

General chair, scientific chair

- Jordan Patracone, Theme Day of GdR IASIS ([Bilevel optimization and hyperparameters learning](#)), 2025, Lyon.
- Amaury Habrard, [Working group GdR IASIS \(Physics aware Machine Learning\)](#), 2025, Paris.

Member of the organizing committees

- Jordan Patracone, [Regional Workshop POPILS](#), 2025, Lyon.

Workshop organization

- Marc Sebban, "Extract full information and meaning from surface imaging" workshop, EUR SLEIGHT Science Event, January 2025, Saint-Etienne.
- Quentin Bertrand, Rémi Emonet, [Generative Models Workshop](#), SMAI 2025, Carcans Maubuisson, 40 people attendance.
- Amaury Habrard Workshop Day, "[Physics Informed Machine Learning and Large Models](#)", Paris, 2025, 100 people attendance.

11.1.2 Scientific events: selection

Member of the conference program committees

- Amaury Habrard, ICML, 2025, Vancouver, area chair.
- Amaury Habrard, NeurIPS, 2025, San Diego, area chair.
- Amaury Habrard, ICLR, 2026, Rio de Janeiro, area chair.
- Rémi Emonet, ICLR Blog post, 2026, Rio de Janeiro, area chair.
- Marc Sebban, SLEIGHT Science Event, Saint-Etienne, 2025, programme committee.

Reviewer

- Quentin Bertrand, ICLR, 2025, Singapour.
- Quentin Bertrand, ICML, 2025, Vancouver.
- Quentin Bertrand, NeurIPS, 2025, San Diego.
- Rémi Emonet, CAp, 2025.
- Rémi Emonet, NeurIPS, 2025.
- Rémi Emonet, AAAI, 2026.
- Rémi Emonet, ICML, 2025.
- Amaury Habrard, CAp, 2025, Strasbourg.
- Jordan Patracone, CAp, 2025, Strasbourg.
- Jordan Patracone, ICML, 2025, Vancouver.
- Jordan Patracone, AAAI, 2025, Singapore.
- Rémi Eyraud, MoL, 2025, New York.
- Eduardo Brandao, AISTATS, 2026, Tangiers.
- Farah Cherfaoui, CAp, 2025.

11.1.3 Journal

Member of the editorial boards

- Amaury Habrard, Journal of Machine Learning Research, editorial board of reviewers.
- Jordan Patracone, Journal of Machine Learning Research.

Reviewer - reviewing activities

- Quentin Bertrand, Journal of Machine Learning Research.
- Quentin Bertrand, Transaction on Machine Learning Research.
- Farah Cherfaoui, journal IEEE Transactions on Information Theory.
- Jordan Patracone, Transactions on Machine Learning Research.
- Jordan Patracone, Journal of Mathematical Imaging and Vision.

11.1.4 Invited talks

- Quentin Bertrand: "*Some Challenges Around Retraining Generative Models on their Own Data*", Imaging in Paris, January 2025.
- Quentin Bertrand, Rémi Emonet: "Generative Models Tutorial", SenHubAI Summer School, April 2025.
- Quentin Bertrand: "*On the Closed Form of Flow Matching: Generalization Does Arise From Target Stochasticity*", NeurIPS, December 2025.
- Rémi Emonet: "*On the Closed Form of Flow Matching: Generalization Does Arise From Target Stochasticity*", EurIPS, December 2025.
- Rémi Emonet "*Evolution of Generative Models: Diving into CFM*", IRISA, Rennes, February 27, 2025,
- Rémi Emonet "*Generalizing and Scaling Optimal Transport*", 10 ans de la FIL, Lyon, May 26, 2025.
- Rémi Emonet "*Generative Models: Diving into CFM*", SenHubIA Spring School, Remote, March 11, 2025.
- Rémi Emonet "*Understanding Generalization In Conditional Flow Matching*", SMAI Minisymposium « Modèles génératifs, OT et restauration d'images », Carcans, June 3, 2025.
- Rémi Emonet "*Creativity in Generative AI*", Design, interfaces and digital perspectives, Saint-Étienne, May 22, 2025.
- Amaury Habrard: "*An introduction to Physics-Informed Machine Learning and Applications*", invited speaker, NormaSTIC Day, April 1, 2025.
- Amaury Habrard: "*What to do with AI in the field of creation—a vision of the limits*". Research Day ECCLA-Esadse-Ensba. January 29, 2025.
- Marc Sebban: "*PDE discovery from scarce data: a key challenge in physics-informed machine learning*", Univ. Alicante, Spain, June 29, 2025.
- Eduardo Brandao "*Measuring the Dynamic Complexity of Self-Organized Systems*", SCCS 2025, May 8, 2025, Istanbul.
- Eduardo Brandao "*A measure of dynamic complexity of self-organized systems*", CCS/FR 2025, June 23, 2025, Paris.

11.1.5 Scientific expertise

- Marc Sebban, scientific expert for the "Dispositif Doctorants" Normandy Region, 2025

11.1.6 Research administration

- Rémi Emonet is head of the Machine Learning project at laboratoire Hubert Curien.
- Amaury Habrard is head of the Data Intelligence team, member of the advisory committee of the Artificial Intelligence Lyon Saint-Etienne consortium AILyS, member of the scientific board of the MIAI AI Cluster, co-head of the AI and ML axis of the *Fédération des Laboratoires d'Informatique de Lyon* (FIL), co-responsible of the working group on Physics-aware Machine Learning of the GdR IASIS
- Marc Sebban is deputy director of the Hubert Curien Lab (UMR CNRS 5516), Member of the Board of Directors of the *Fédération des Laboratoires d'informatique de Lyon* (FIL), Member of the executive committee of the EUR Manutech Sleight, Member of the COMEX Labex MILYON, Member of the COS Inria Lyon Centre.

11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

11.2.1 Teaching

- Quentin Bertrand:
 - Computer Science Master, Deep Generative Models, 10h, M2, ENS Lyon.
 - Computer Science Master, Numerical Optimal Transport for Machine and Deep Learning, 10h, M2, ENS Lyon.
- Eduardo Brandao:
 - Diplôme Universitaire Cycle initial en Technologie de l'Information de Saint-Étienne: Mathématiques pour l'ingénieur 1, 48h, L1, Télécom Saint-Étienne, UJM
 - Diplôme Universitaire Cycle initial en Technologie de l'Information de Saint-Étienne: Mathématiques pour l'ingénieur 2, 49.5h, L1, Télécom Saint-Étienne, UJM
 - Formation Ingénieur en Apprentissage Data Engineering: Big Data Project, 11h, 3A, Télécom Saint-Étienne, UJM
- Farah Cherfaoui:
 - Licence Mathématiques-Physique-Chimie: Outils Informatiques, 83h, L1, Faculté des Sciences, UJM.
 - Licence informatique: Programmation Impérative, 72h, L1, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Advanced Machine Learning, 4h, M2, Faculté des Sciences, UJM.
 - Licence 3 informatique: Probabilités-Statistiques, 24h, L3, Faculté des Sciences, UJM.
 - Licence 3 informatique: Langage formel 2, 36h, L3, Faculté des Sciences, UJM.
- Rémi Emonet:
 - Master Machine Learning and Data Mining: Probabilistic Graphical Models, 20h, M2, Faculté des Sciences, UJM.
 - Master Données et Systèmes Connectés: Programmation Web Avancée, 30h, M1, Faculté des Sciences, UJM.
- Rémi Eyraud:
 - Master Machine Learning and Data Mining: Research Methodology, 20h, 1A, Faculté des Sciences, UJM.
 - Master Données et Systèmes Connectés: Research Methodology, 20h, 2A, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Machine Learning Fundamentals and Algorithms, 30h, 1A, Faculté des Sciences, UJM.
 - Master Données et Systèmes Connectés: Machine Learning Fundamentals and Algorithms, 30h, 1A, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Advanced Machine Learning, 3h, M2, Faculté des Sciences, UJM.
 - GACO: Base de Données, 18h, 2A, IUT de Saint-Etienne, UJM.
 - GACO: Conception de Sites Web Dynamiques, 112h, 2A, IUT de Saint-Etienne, UJM.
 - GACO: Etablir le Diagnostic Marketing d'une Organisation (SAE), 8h, 2A, IUT de Saint-Etienne, UJM.
 - GACO: Traitement des Données, 20h, 3A, IUT de Saint-Etienne, UJM.
 - GACO: Environnement Informatique, 16h, 1A, IUT de Saint-Etienne, UJM.

- Benjamin Girault:
 - Master Machine Learning and Data Mining: Deep Learning I, 20h, M1, Faculté des Sciences, UJM.
- Amaury Habrard:
 - Master Machine Learning and Data Mining: Advanced Algorithms and Programming, 20h, M2, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Advanced Machine Learning, 20h, M2, Faculté des Sciences, UJM.
 - joint course Master Machine Learning and Data Mining and Master Données and Systèmes Connectés: Deep Learning II, 15h, M2, Faculté des Sciences, UJM.
- Jordan Patracone:
 - Formation Ingénieur Télécom Saint-Étienne: Langage C Algorithmie et structures de données, 13.5h, 1A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Algorithmique et structures de données, 4.5h, 1A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Projet recherche et innovation, 5h, 3A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Big Data Project, 10h, 3A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Projet d'ingénierie, 14h, 2A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Python et sciences des données, 18h, 2A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur Télécom Saint-Étienne: Machine learning, 37.5h, 2A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur en Apprentissage Data Engineering: Statistiques inférentielles, 33h, 2A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur en Apprentissage Data Engineering: Algorithms for data analysis, 14h, 2A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur en Apprentissage Data Engineering: Qualité de fonctionnement, 11h, 3A, Telecom Saint-Etienne, UJM.
 - Formation Ingénieur en Apprentissage Data Engineering: Innovation en data, 2h, 3A, Telecom Saint-Etienne, UJM.
- Marc Sebban:
 - Licence Informatique: Probabilités-Statistiques, 24h, L3, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Introduction to Machine Learning, 20h, M1, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Data Analysis, 24h, M1, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Advanced Machine Learning, 12h, M2, Faculté des Sciences, UJM.
 - Master Machine Learning and Data Mining: Machine Learning Project, 20h, M2, Faculté des Sciences, UJM.

11.2.2 Supervision

- PhD in progress: Michael Gault. Développement et évaluation d’approches hybrides de machine learning informé par la physique pour la modélisation d’une unité de captage de CO₂, since Octo. 2025. Marc Sebban and Rémi Emonet.
- PhD in progress: Petros Kafkas. Harmonic Neural Networks for Physics-informed Machine Learning, since Octo. 2025. Marc Sebban and Benjamin Girault.
- PhD in progress: Diego Pinto-Suárez, Monotone Operators and Neural Architectures: Leveraging Interactions for physically Structured Approximations, since Dec. 2025. Jordan Patracone, Marc Sebban.
- PhD defended in 12/2025: Fayad Ali Banna. From Energy Absorption to Self-Organization: A Physics-Informed Machine Learning Approach to Laser-Matter Interaction. Marc Sebban and Rémi Emonet.
- PhD in progress: Hind Atbir. Learning fair and robust kernel-based models with generalization guarantees, since Oct. 2024. Farah Cherfaoui.
- Postdoc in progress: Antoine Caradot. On the estimation of integrals of functions: applications in PIML, since Oct. 2024. Rémi Emonet, Amaury Habrard and Marc Sebban.
- PhD in progress: Wissal Ghamour. Nouvelles approches de l’Intelligence Artificielle pour les attaques par canaux auxiliaires, since Dec. 2024. Rémi Eyraud.
- PhD defended in 12/2025: Dorian Llavata. Apprentissage Profond Diversement Supervisé Pour Les Attaques Par Canaux Auxiliaires, Rémi Eyraud.
- PhD in progress: Robin Mermillod-blondin, Optimisation multicritère de couleurs plasmoniques pour les documents d’identité. Rémi Emonet.
- PhD defended in 11/2025: Sayan Chaki, Unsupervised Analysis of Ornaments extracted from Ancient Books, Rémi Emonet.
- PhD in progress: Abdel-Rahim Mezidi, Unveiling and Incorporating Knowledge in Physics-Guided Machine Learning Models, Amaury Habrard, Jordan Patracone.
- PhD in progress: Ben Gao, Toward frugal machine learning with physics-aware models, Jordan Patracone.
- PhD in progress: Thibault Girardin, Prediction of multidimensional colors printed by laser on plasmonic metamaterials using deep learning and adaptive strategies, Amaury Habrard.
- PhD in progress: Erick Gomez, TRansfer lEarning for Frugal and Accurate modeling of SURface Functionalization prediction – application to multicomponent alloys, Amaury Habrard.
- PhD Defended in 01/2025: Volodimir Mitarchuk, Theory and Algorithms for the Understanding of Deep Neural Network on Sequential Data, Rémi Eyraud, Amaury Habrard, Rémi Emonet. adaptive strategies.
- Postdoc (Jan-Oct 2025): Volodimir Mitarchuk, Operator learning in reproducing kernel Banach spaces, Jordan Patracone.

11.2.3 Juries

- Marc Sebban: Fabien Lioni (PhD, Univ. Côte d’Azur, reviewer), Alice Lacan (PhD, Univ. Paris-Saclay, President), Lawrence Stewart (PhD, Univ. Paris-Saclay, President).
- Rémi Emonet: Mehran Adibi (PhD, Université de Lorraine, Reviewer), Fayad Ali Banna (PhD, Univ. Saint-Etienne, Supervisor), Sayan Chaki (PhD, Univ. Saint-Etienne, Supervisor).

- Rémi Eyraud: Chandrasekar Subamani Narayana (PhD, Aix-Marseille Univ., reviewer), Volodimir Mitarchuck (PhD, Univ. Saint-Etienne, Director)
- Amaury Habrard: Myriam Tami (HDR, Centrale Supélec-Univ. Paris-Saclay, Reviewer), Marc Lafon (PhD, CNAM - Sorbonne Univ, Reviewer), Jean-Rémy Conti (PhD, Telecom Paris, Reviewer), Rui Yang (PhD, Ecole Centrale de Lyon, Member), Louis Serrano (PhD, Sorbonne Univ., Reviewer) Lucas Gnecco-Heredia (PhD, Univ Paris-Dauphine PSL, member), Kanishka Gosh Dastidar (PhD, Univ. Passau, Germany, Reviewer), Volodimir Mitarchuk (PhD, Univ. Saint-Etienne, Co-director)

11.3 Popularization

11.3.1 Productions (articles, videos, podcasts, serious games, ...)

- Rémi Emonet, Quentin Bertrand, collaboration with Ockham, we wrote a [friendly blog post](#) on recent flow matching techniques, April 2025.
- Rémi Eyraud: [filmed interview](#) at the radio Ici Saint-Etienne during the AI international submit in Paris, 10/02/2025.
- Quentin Bertrand: our recent works on self-consuming generative models and their biases got media coverage from [Sciences et Avenir!](#) 09/2025.

11.3.2 Participation in Live events

- Rémi Eyraud: 45 min of conference titled "Qu'est-ce que l'IA?", festival du cinéma jeune public *Tête de mule*, Cinéma Méliès de Saint-Etienne, 200 kids, 27/04/2025.
- Rémi Eyraud: 2h of conference titled "de l'Algorithme à l'IA". Complètement conf³ – Mairie de Saint-Genest-Lerpt, 50 persons, 13/05/2025.
- Amaury Habrard and Rémi Eyraud: an evening of science popularization titled "[Vous ne direz plus Intelligence Artificielle !](#)", festival Pint of Science, pub 6 Nations, Saint-Etienne, 50 persons, 21/05/2025
- Rémi Eyraud: 2h of conference and workshops titled "Du dieu Algorithme à la déesse IA", Service Universitaire de Pédagogie, IUT de Saint-Etienne, 28/05/2025.
- Rémi Eyraud: 2h of conference and workshops titled "L'IA et les Associations Familiales", Union des Associations Familiales de la Loire, Chateau de Magneux-Haute-Rive, 70 persons, 10/10/2025.
- Rémi Eyraud: 2h of conference titled "Démystifier l'IA". L'escale, Saint-Martin-en-Haut, 3/11/2025.
- Rémi Eyraud: 1h30 of conference titled "Liberté, Algorithmes et Responsabilités face à l'IA", semaine de la laïcité, Médiathèque de Saint-Etienne, 13/12/2025.

11.3.3 Others science outreach relevant activities

- Amaury Habrard: round table "Que faire de l'IA dans le champ de la création?", Journée d'étude ECCLA-Esadse, 28/01/2025.
- Rémi Eyraud: 30 minutes of conference followed by 3h of workshop on "Centres sociaux et IA", centres sociaux Loire et Haute-Loire, 70 persons, 8/04/2025.
- Rémi Eyraud: 2h of workshop titled "Explorons l'IA", Fête de la science, Explora Saint-Etienne, collégiens de 4ème et 3ème, 30 teens, 10/10/2025.

12 Scientific production

12.1 Major publications

- [1] Q. Bertrand, J. A. Duque, E. Calvano and G. Gidel. ‘Self-Play Q-Learners Can Provably Collude in the Iterated Prisoner’s Dilemma’. In: *Proceedings of the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025*. International Conference on Machine Learning. Vancouver (BC), Canada, 14th July 2025. URL: <https://hal.science/hal-05120296> (cit. on p. 8).
- [2] Q. Bertrand, A. Gagneux, M. Massias and R. Emonet. ‘On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity’. In: *NeurIPS 2025. NeurIPS 2025 - 39th Annual Conference on Neural Information Processing Systems*. San Diego (CA), United States, 1st Dec. 2025. URL: <https://hal.science/hal-05119547> (cit. on pp. 8, 11).
- [3] A. Caradot, R. Emonet, A. Habrard, A.-R. Mezidi and M. Sebban. ‘Provably Accurate Adaptive Sampling for Collocation Points in Physics-informed Neural Networks’. In: *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases*. Vol. 16017. Lecture Notes in Computer Science. Porto, Portugal: Springer Nature Switzerland, 25th Sept. 2025, pp. 19–37. DOI: [10.1007/978-3-032-06096-9_2](https://doi.org/10.1007/978-3-032-06096-9_2). URL: <https://hal.science/hal-05017201> (cit. on p. 12).
- [4] A. Gagneux, S. Martin, R. Emonet, Q. Bertrand and M. Massias. *A Visual Dive into Conditional Flow Matching*. Apr. 2025. URL: <https://hal.science/hal-05088592> (cit. on p. 11).
- [5] B. Gao, J. Patracone, S. Chrétien and O. Alata. ‘Conformal Online Learning of Deep Koopman Linear Embeddings’. In: *In Advances in Neural Information Processing Systems. NeurIPS 2025 - 39th Annual Conference on Neural Information Processing Systems*. San Diego (California), United States, 2025. URL: <https://inria.hal.science/hal-05086738> (cit. on pp. 8, 12).
- [6] A.-R. Mezidi, J. Patracone, S. Salzo, A. Habrard, M. Pontil, R. Emonet and M. Sebban. ‘A Bregman Proximal Viewpoint on Neural Operators’. In: *International Conference on Machine Learning, Vancouver, Canada, 2025*. URL: <https://inria.hal.science/hal-04584456> (cit. on pp. 8, 12).

12.2 Publications of the year

International journals

- [7] G. Bonafos, F. Vial, C. Cornec, L. Papet, L. Lego, J. Audouard, A. de Vito, D. Reby, J. Rouch, R. Emonet, H. Patural and N. Mathevon. ‘Premature baby cries remain different from those of full-term babies’. In: *Journal of the Acoustical Society of America* 157.6 (2025), pp. 4175–4183. DOI: [10.1121/10.0036842](https://doi.org/10.1121/10.0036842). URL: <https://hal.science/hal-05245427>.
- [8] D. Llavata, E. Cagli, R. Eyraud, V. Grosso and L. Bossuet. ‘Unsupervised horizontal attacks against public-key primitives with DCCA’. In: *IACR Communications in Cryptology* 2.1 (8th Apr. 2025). DOI: [10.62056/ah5w7ta5v](https://doi.org/10.62056/ah5w7ta5v). URL: <https://hal.science/hal-05266270>.
- [9] R. Mermillod-Blondin, N. Dalloz, A. Tremeau, A. W. Fadhilah, R. Emonet and N. Destouches. ‘Image quality metrics for restricted gamut images produced by laser-induced printing on plasmonic films’. In: *Journal of the European Optical Society : Rapid publications* 21.2 (2025), p. 36. DOI: [10.1051/jeos/2025025](https://doi.org/10.1051/jeos/2025025). URL: <https://hal.science/hal-05204435> (cit. on p. 14).
- [10] L. Minier, J. Rouch, B. Sabbagh, F. Bertucci, E. Parmentier, D. Lecchini, F. Sèbe, N. Mathevon and R. Emonet. ‘Visualization and quantification of coral reef soundscapes using CoralSoundExplorer software’. In: *PLoS Computational Biology* 21.4 (10th Apr. 2025), e1012050. DOI: [10.1371/journal.pcbi.1012050](https://doi.org/10.1371/journal.pcbi.1012050). URL: <https://hal.science/hal-05034375>.
- [11] S. Venturini, M. de Santis, J. Patracone, F. Rinaldi, S. Salzo and M. Schmidt. ‘Relax and penalize: a new bilevel approach to mixed-binary hyperparameter optimization’. In: *Transactions on Machine Learning Research Journal* (Feb. 2025). URL: <https://hal.science/hal-04183917> (cit. on p. 15).

International peer-reviewed conferences

- [12] Q. Bertrand, J. A. Duque, E. Calvano and G. Gidel. ‘Self-Play Q-Learners Can Provably Collude in the Iterated Prisoner’s Dilemma’. In: *Proceedings of the 42nd International Conference on Machine Learning, Vancouver, Canada. PMLR 267, 2025*. International Conference on Machine Learning. Vancouver (BC), Canada, 14th July 2025. URL: <https://hal.science/hal-05120296> (cit. on p. 15).
- [13] Q. Bertrand, A. Gagneux, M. Massias and R. Emonet. ‘On the Closed-Form of Flow Matching: Generalization Does Not Arise from Target Stochasticity’. In: *NeurIPS 2025. NeurIPS 2025 - 39th Annual Conference on Neural Information Processing Systems. San Diego (CA), United States, 1st Dec. 2025*. URL: <https://hal.science/hal-05119547>.
- [14] A. Caradot, R. Emonet, A.-R. Mezidi, A. Habrard and M. Sebban. ‘Méthode de quadrature pour les PINNs fondée théoriquement sur la hessienne des résiduels’. In: 30ème Colloque sur le traitement du signal et des images GRETSI. Strasbourg, France, Aug. 2025. URL: <https://hal.science/hal-05128227>.
- [15] S. Chrétien, B. Gao, J. Patracone and O. Alata. ‘Using Signatures and Koopman Operator to Learn Nonlinear Dynamics’. In: *7th International Conference of Geometric Science of Information - GSI 2025. GSI’25 - 7th International Conference on Geometric Science of Information. saint-malo, France, 2025*. URL: <https://inria.hal.science/hal-05224359>.
- [16] B. Gao, J. Patracone, S. Chrétien and O. Alata. ‘Conformal Online Learning of Deep Koopman Linear Embeddings’. In: *In Advances in Neural Information Processing Systems. NeurIPS 2025 - 39th Annual Conference on Neural Information Processing Systems. San Diego (California), United States, 2025*. URL: <https://inria.hal.science/hal-05086738>.
- [17] B. Gao, J. Patracone, S. Chrétien and O. Alata. ‘When to Learn: Conformal Scores as Online Update Criteria’. In: Workshop on Predictions and Uncertainty at COLT 2025. Lyon, France, 30th June 2025. URL: <https://inria.hal.science/hal-05224472>.
- [18] M. Gault, P. Bachaud, B. Celse, R. Emonet and M. Sebban. ‘Physics-Informed Machine Learning for Modeling CO2 Capture from Scarce Data’. In: 37th IEEE International Conference on Tools with Artificial Intelligence. 2025 IEEE 37th International Conference on Tools with Artificial Intelligence (ICTAI). Athènes, Greece: IEEE, 5th Nov. 2025, pp. 1436–1442. doi: [10.1109/ICTAI66417.2025.00207](https://doi.org/10.1109/ICTAI66417.2025.00207). URL: <https://hal.science/hal-05313730> (cit. on p. 13).
- [19] A.-R. Mezidi, J. Patracone and A. Habrard. ‘Douglas-Rachford Splitting for Hybrid Differentiable Models’. In: Workshop on Differentiable Systems and Scientific Machine Learning (DiffSys). Copenhagen, Denmark, 6th Dec. 2025. URL: <https://inria.hal.science/hal-05343823>.
- [20] A.-R. Mezidi, J. Patracone, S. Salzo, A. Habrard, M. Pontil, R. Emonet and M. Sebban. ‘A Bregman Proximal Viewpoint on Neural Operators’. In: International Conference on Machine Learning. Vancouver, Canada, 2025. URL: <https://inria.hal.science/hal-04584456>.

Conferences without proceedings

- [21] F. Ali Banna, E. Brandão, A. Nakhoul, R. Emonet, M. Sebban and J.-P. Colombier. ‘Learning dynamics in ultrafast laser self-organization’. In: Progress in Ultrafast Laser Modifications of Materials. Villars-sur-Ollon, France, 29th June 2025. URL: <https://ujm.hal.science/ujm-05214760> (cit. on p. 14).
- [22] G. Bonafos, J. Rouch, L. Lego, D. Reby, H. Patural, N. Mathevon and R. Emonet. ‘Speech transformer models for extracting information from baby cries’. In: Workshop on Child Computer Interaction - WOCCI 2025. Nijmegen, Netherlands: ISCA, 22nd Aug. 2025, pp. 11–15. doi: [10.21437/wocci.2025-3](https://doi.org/10.21437/wocci.2025-3). URL: <https://hal.science/hal-05304733>.
- [23] A. Caradot, R. Emonet, A. Habrard, A.-R. Mezidi and M. Sebban. ‘Provably Accurate Adaptive Sampling for Collocation Points in Physics-informed Neural Networks’. In: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. Vol. 16017. Lecture Notes in Computer Science. Porto, Portugal: Springer Nature Switzerland, 25th Sept. 2025, pp. 19–37. doi: [10.1007/978-3-032-06096-9_2](https://doi.org/10.1007/978-3-032-06096-9_2). URL: <https://hal.science/hal-05017201>.

- [24] B. Gao, J. Patracone, O. Alata and S. Chrétien. ‘Apprentissage incrémental de l’opérateur de Koopman pour systèmes non-autonomes via prédiction conforme’. In: Colloque GRETSI’25. Strasbourg, France, Aug. 2025. URL: <https://inria.hal.science/hal-05024738>.
- [25] T. Girardin, N. Destouches and A. Habrard. ‘Contextual Hypernetwork for Adaptive Prediction of Laser-Induced Colors on Quasi-Random Plasmonic Metasurfaces’. In: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases. Porto (Portugal), France, 15th Sept. 2025. URL: <https://hal.science/hal-05271900> (cit. on p. 13).

Reports & preprints

- [26] F. Ali Banna, E. Brandão, A. Nakhoul, R. Emonet, M. Sebban and J.-P. Colombier. *Photonic Self-Learning in Ultrafast Laser-Induced Complexity*. 28th Mar. 2025. URL: <https://hal.science/hal-05012225> (cit. on p. 14).
- [27] F. Ali Banna, A. Caradot, E. Brandão, J.-P. Colombier, R. Emonet and M. Sebban. *Unrolled-SINDy: A Stable Explicit Method for Non linear PDE Discovery from Sparsely Sampled Data*. 20th Oct. 2025. URL: <https://inria.hal.science/hal-05324777>.
- [28] H. Atbir, F. Cherfaoui, G. Metzler, E. Morvant and P. Viallard. *PAC-Bayesian Bounds on Constrained f -Entropic Risk Measures*. Oct. 2025. URL: <https://hal.science/hal-05314302>.
- [29] A. Caradot and Z. Lin. *Differential graded algebras with divided powers and homotopy Lie algebras*. 27th Nov. 2025. URL: <https://inria.hal.science/hal-05386156>.
- [30] S. Chrétien, B. Gao, J. Patracone, A.-R. Mezidi and O. Alata. *Tukey-Median of Means-Gradients for Langevin Dynamics: a Robust Approach to Fitting Machine Learning Models*. 17th Oct. 2025. URL: <https://inria.hal.science/hal-05319256>.
- [31] T. Feng, B. M. Booth, K. Mundnich, E. Zhou, B. Girault, K. Lerman and S. Narayanan. *TILES-2018 Sleep Benchmark Dataset: A Longitudinal Wearable Sleep Data Set of Hospital Workers for Modeling and Understanding Sleep Behaviors*. 4th July 2025. URL: <https://inria.hal.science/hal-05302189>.

Other scientific publications

- [32] A. Gagneux, S. Martin, R. Emonet, Q. Bertrand and M. Massias. *A Visual Dive into Conditional Flow Matching*. Apr. 2025. URL: <https://hal.science/hal-05088592>.