Hi! PARIS

PARIS ARTIFICIAL INTELLIGENCE FOR SOCIETY

EMPOWERING SOCIETY WITH DATA & AI

Founding members

THE **CENTER**

Hi! PARIS is the result of a unique collaboration between Institut Polytechnique de Paris and HEC Paris. It is a first-class hub for France and global corporate partners who reap the benefits of working closely with our rich scientific talent pool and students.

Based on joint expertise and a multidisciplinary approach, the Center addresses key challenges related to technological transformation and its impact on business and society. Theoretical and methodological research in AI and Data analytics is conducted at the highest level. The Center tackles the managerial, legal, economic, ethical and societal issues emerging due to exponentially larger data sets harnessed through artificial intelligence.

Ambition

The global ambition of this new interdisciplinary center is to ensure that AI and data empower business and society. It will provide a unique framework for research, education (engineers, managers, young researchers, executive education), innovation, and technology transfer to businesses. It will take advantage of cross-fertilization between fundamental sciences, technology, management and social sciences, all of which are fields of excellence for both Institut Polytechnique de Paris and HEC Paris. These resources are at present essential for companies and laboratories, both public and private.

It also aims at stimulating productive interactions between researchers, students and organizations, thus enabling the emergence of high-potential projects, up to startups. The Center's ambition as regards AI and Data Analytics is to compete with the very best international institutions.

Hi! PARIS is a destination of choice for the most talented students and researchers from all over the world, all of whom address questions related to data science, artificial intelligence, their role in science, technology and business, and impact on society.

Hi! PARIS exceptional growth directly impacts the success of Paris and France's global leadership in AI. By attracting international talent, **Hi! PARIS** has an economic, social, and scientific impact that strengthens France and Europe's leadership positions.

Foreword

Gaël Richard
Executive director
Professor at Télécom Paris

June 2022

Hi! PARIS is an interdisciplinary center for Data Analytics and Artificial Intelligence for Science, Business and Society created by Institut Polytechnique de Paris (IP Paris) and HEC Paris and recently joined by Inria Saclay.

Hi! PARIS conducts frontier multidisciplinary research from fundamental problems on methods for AI and data analytics to business applications across all sectors and covering implications for society.

The research domains covered by **Hi! PARIS** include in particular:

- Foundations of AI and Machine Learning (reinforcement learning, deep learning, mathematics for AI, natural language processing, federated learning, symbolic AI, interpretable and explainable AI, cybersecurity, data, etc.),
- AI and Datascience for Business (digital transformation, AI for marketing or finance, data economy, etc.),
- and AI and Data Science for Society (sustainable AI, frugal AI, AI for climate, ethics and bias in AI, robotics, neurosciences, AI and justice, AI and health, privacy, etc.).

The goal of this document is not to exhaustively describe this research in all its diversity but rather to provide *some visions of research in AI and Data Analytics* with a specific focus on some of the emblematic projects of the **Hi! PARIS** chair holders and fellows, published in major peer-reviewed journals and international conference proceedings. This document thus shines a spotlight on specific research topics that all represent major challenges for the future and that hold great promise.

I wish you a pleasant reading experience and hope you will find food for thought and ideas for further collaboration in these pages.

Gaël Richard is Grand prix IMT-National Academy of Science 2020, and IEEE Fellow.

Table of contents

Visions of research

Eric Moulines
Scientific co-director
Professor at École Polytechnique

Nicolas Vieille
Scientific co-director
Professor at HEC Paris

Fantastic advances in machine learning (ML) and artificial intelligence (AI) over the past decade have generated tremendous interest from the scientific community, industry, governments, and society at large. There has been a flurry of research initiatives, a flood of investment, and thriving startup activity to capitalize on these advances and further develop the field.

Historically, the field of AI has experienced several peaks of inflated expectations followed by disillusionment—for example, in the 1970s and 1980s—leading to the “AI winter” that saw massive funding cuts, limited industry adoption, and the end of targeted research activity in the field. But there is no reason to fear that the recent upsurge in interest in AI is another exaggerated phenomenon that will soon fizzle out. Unlike other times, AI has already opened up many exciting research avenues, added a lot of value and become a central part of numerous product strategies. It will continue to do so.

Artificial Neural Networks, a simple, still promising idea

Machine learning has been a very active research topic since the 1970s. After more or less failed attempts to use symbolic approaches and to model reasoning using first-order logic, statistical approaches have been used since the mid-1980s. Bayesian networks, decision trees, and support vector machines have long been at the forefront of such approaches. But none of these methods has been as successful as **Deep Learning**, which has spurred the current surge in AI.

Initial research in neural networks was motivated by the observation that “human intelligence” appears to arise from highly interconnected networks of simple, nonlinear “neurons” that learn by adjusting the strength of their synaptic connections. This observation led to the central scientific question of how such networks of elementary computational units can learn the complicated internal representations required for extremely demanding tasks such as visual scene recognition, speech recognition, or speech-to-speech translation. Deep Learning attempts to answer this question by using many layers of “formal neurons” as representations and learning the synaptic weights and biases that parameterize these neurons by following the stochastic gradient of an objective function that measures how well the network performs.

> *Neural networks played a key role in modern AI*

It may seem surprising that such a conceptually simple computational model has proven so effective when applied to large training sets! But the truth is that the neural networks of yesterday and the deep networks of today have played a key role in the tremendous advances in AI.

Eric Moulines is CNRS Silver medal 2010, ORANGE Grand Prix of Science Academy 2011, Elected member of the French Academy of Sciences 2017, Elected member of the Institute of Mathematical Statistics (IMS) 2016.

Nicolas Vieille is Lanchester Prize 2003 in Operations Research, fellow of the Econometric Society, fellow of the Game Theory Society.

Remarkable achievements...

Deep Learning reinvigorated neural network research in the mid-2000s by introducing several elements that facilitated the training of deeper networks. The advent of graphics processing units (GPUs) and the availability of large datasets were important enablers of Deep Learning and were greatly enhanced by the development of flexible open-source software platforms including, alongside many other goodies, automatic differentiation, which made it possible to consider far more complex models with minimal development effort. The availability of these tools made it easy to train extremely complex deep networks, with hundreds of millions of parameters. Software platforms also allowed developers to reuse the latest models and their building blocks. Perhaps surprisingly, it was the composition of more layers and apparently "minor" tweaks that enabled more complex nonlinearities and have broken through problems that seemed out of reach.

Deep Learning has already outperformed several benchmarks in computer vision, speech recognition, text-to-speech synthesis, image generation, reinforcement learning and machine translation. Deep neural networks are now tackling new challenges in fields as varied as drug discovery, particle accelerator data analysis, genetics and neuroscience.

> *We are far from the end of the AI adventure!*

...and numerous challenges still to be tackled

Despite the great advances in the field of artificial intelligence and Deep Learning, current methods are not devoid of shortcomings, some of which are still prohibitively expensive. We are therefore far from the end of the AI adventure! While Deep Learning excels in benchmarks for natural language processing and computer vision, it often under-performs in real-world applications—any interaction with a chatbot will remind us how far we still have to go. Deep Learning has limited ability to transfer knowledge, does not adapt quickly to changing tasks or distributions, and insufficiently incorporates world or prior knowledge.

In addition, there are numerous concerns about the black-box nature of Deep Learning algorithms and the trustworthiness of solutions. Deep Learning models are also highly susceptible to adversarial examples that are barely perceptible to humans but can easily fool the ML model and lead to misclassification. Such erroneous results can be catastrophic in a safety-critical technical environment with long-term financial and legal implications. Indeed, Deep Learning models have been shown to fail on new data, new applications, deployments in the wild, and stress testing.

Finally, Deep Learning in its current form is data hungry and computationally intensive. Recent estimates suggest that further increases in the power of such systems are economically, technically, and environmentally unsustainable.

Hi! PARIS is the research hub in this area. Our goal is to overcome the current limitations of AI, close the gap with real-world applications, and contribute to the AI adventure! We have an exceptional community of scholars and researchers at **Hi! PARIS** addressing these challenges, and believe us, they are numerous and particularly diverse.

Artificial Intelligence more and more deeply embedded in organizations…

Current progress in AI, and especially in prediction technology, is impressive. ML tools are becoming easier to use and can be deployed by non-experts. As a general-purpose technology, AI has a potentially disruptive impact on organizations, firms and society at large.

This shift is underway. Algorithms are already responsible for a large majority of transactions in financial markets, through algorithmic or high-frequency trading. AI is routinely used to provide financial services, or to retrieve information from unstructured data to help with forecasting and nowcasting, in finance and accounting. In operations, AI is used to handle inventories and logistics. E-commerce, online platforms, but also energy providers increasingly rely on such tools to help foster their efficiency, by reducing search frictions for consumers or improving demand predictions. AI products can increasingly be used to write reports or translate technical documents, for example. They also contribute to a new innovation ecosystem by providing solutions for outsourcing many business processes or optimizing product design.

At the societal level, the opposition between a gloomy view (with algorithms replacing humans in a seemingly unlimited set of tasks, creating a powerful displacement effect) and a rosy assessment (with technological shifts leading to increased aggregate prosperity) is misleading. On the one hand, new tasks appear, involving complex reasoning or abstract problem-solving skills. And sometimes, the demand for "old" tasks may remain: As an historical example, the introduction of ATMs, far from putting bank tellers out of their job, has in fact increased their number, by enabling them to focus on more productive tasks. At the same time though, costly changes in business processes are needed, and it is widely asserted that this may lead to increasing inequality in the short- to mid-term due to skill bias, and to an increased share of capital in the economy. To some extent, country-level rents may be expected, creating challenging HR management and public policy intervention questions.

> *AI is routinely used to provide services in organizations*

…raising recurrent issues on privacy, fairness, transparency…

AI solutions ultimately rely on data coming from humans, raising many issues. There is a clear demand for privacy by consumers and citizens: The personalization of products (using taste information derived from AI) is acceptable only to the extent that it comes together with a sense of control over the use of this data. Yet, this demand is complex. It depends on age, on context, and it is prone to a number of behavioral effects, well-documented in the academic literature, that affect how agents make decisions over data privacy issues. This in turn paves the way to potential nudges, by private firms or public entities.

Imagine the following scenario. A few years ago, Mr. A submitted genetic information to, say a genealogy site X. A consent form was signed by A, which however did not anticipate the fact that X would be bought out by an insurance company Y. Mr. B, a sibling of A, is now submitting a loan request, that is screened by Y. The decision whether to grant the loan (or the price of health insurance) is based on health predictions for the duration of the loan, using AI prediction tools that may potentially use social network data to infer that A and B are siblings, and genetic data from A. Such an admittedly contrived scenario highlights the durability of the information contained in some data, possibly unanticipated uses of data (both genetic and unstructured social network data), spillover effects on B of data casually shared by A, and a blurred distinction between the private and public spheres.

Such issues clearly raise regulatory and legal challenges. One aspect of a legal response consists in designing an appropriate regulatory framework, thus requiring an in-depth assessment of the ethical risks, and an examination of the trade-offs inherent to policy decisions involving ethical issues. Beyond privacy, this also applies to fairness and non-discrimination. For instance, there are competing notions of non-discrimination, and imposing it on the outcomes of an algorithm, or on the algorithm itself, is not equivalent. Legal and corporate responses thus far, such as the General Data Protection Regulation (GDPR), the planned removal of third-party cookies, or the discussions around "the need for an explanation", have an obvious impact on business practices, especially for marketing and human resources.

> *All these challenges need a pluri-disciplinary approach*

Whatever the legal framework, an additional requirement will be the development of algorithmic legal tools to check compliance, e.g., to make sure that the algorithms implemented do comply with differential privacy requirements that may be imposed, or, in another domain, to check that the algorithms used in banking for stress tests are consistent with the regulation.

The increased use of algorithms for policy decisions leads to new research questions, and to more pluri-disciplinary work involving social and computer scientists, and engineers. To limit ourselves to just a few examples, deciding in a healthcare context which patients should receive a treatment is a causal inference problem, not a pure prediction problem, and calls for methods combining the strengths of ML and those of more traditional statistics. As another illustration, fascinating questions emerge when addressing the fact that algorithms will interact more and more, and need to address the fact that the data from which they learn are endogenous. What is the potential for competitors to achieve collusive outcomes using pricing algorithms? How will the software of an autonomous car perform when surrounded with other autonomous cars? How do the interactions with machines and/or algorithms impact the way agents make decisions? All these questions and many others to come clearly show to what extent social and technological considerations are entangled and demand a combined approach.

Whereas academics tend to be unambiguously optimistic about the long-term potential productivity gains from AI, these gains have only partially materialized so far, following a pattern observed in earlier paradigmatic shifts (computers and Internet, for the most recent).

The main reason for that is a lack of skills, which points to the need to expand the training of people combining business knowledge with technical expertise—one of the chief goals of the **Hi! PARIS** center.

Changing engineering disciplines

Integrating domain knowledge models and data-driven approaches

One of the major novel approaches followed at **Hi! PARIS** for elaborating future AI solutions, and a topic Eric Moulines is specifically contributing to, is **Data-Centric Engineering** (DCE). DCE is about leveraging domain-specific knowledge and integrating mechanistic models or other forms of symbolic thinking with data-driven processing. There has been some initial success in applying ML /AI models in a “plug-and-play” manner, without the need to change the underlying algorithms for domain-specific applications. However, to solve the more challenging problems in each discipline, ML /AI algorithms must be adapted to incorporate domain knowledge in addition to data-driven methods. In many domains, especially in the design phase of prototypes, very little data is available. Domain knowledge-driven models are more useful in such situations than Deep Learning methods, which require huge data collections. Data-driven approaches do not encode physical laws such as the conservation of mass, momentum, or energy that form the basis of most engineering applications: incorporating the knowledge acquired by scientists and engineers since the beginning of mankind allows us to considerably reduce the amount of data needed!

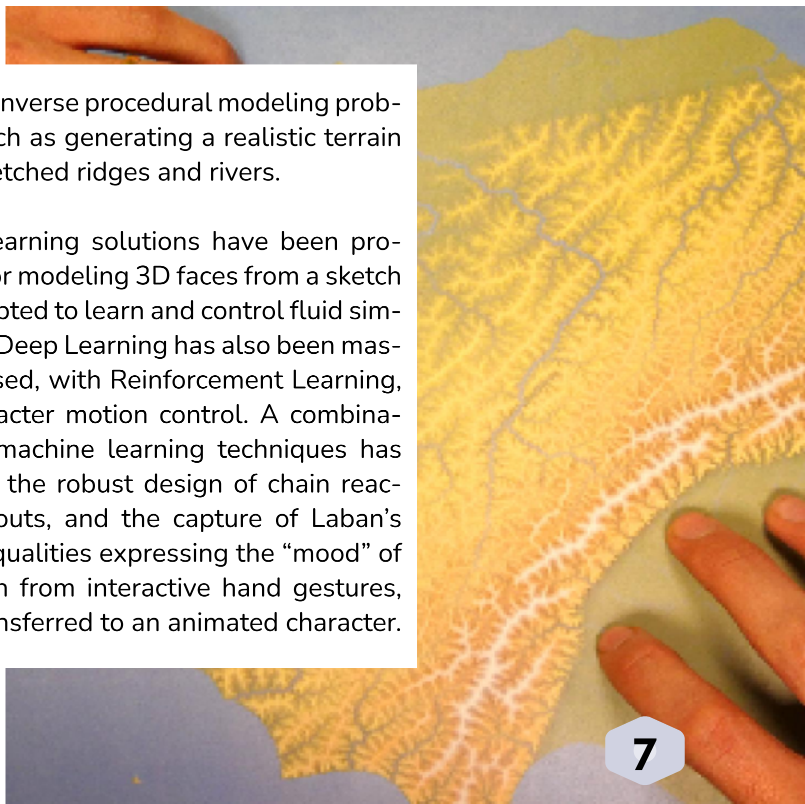
The convergence of simulations, ML, and statistical algorithms, combined with hardware improvements such as high-performance graphics processing units (GPUs), high computing power, cheap streaming sensors, and inexpensive storage, fundamentally change traditional engineering disciplines. In the engineering design phase, improvements such as faster product prototyping, shorter time-to-market, the ability to algorithmically generate and explore different design spaces and solutions, and data- and simulation-driven “what-if” scenarios for effective decision-making will all add value. In the technical operations phase, enhancements will include integrated simulations and data-driven solutions for improved process optimization, quantitative reliability and risk assessments, operations planning, and scheduling, plant monitoring and outage prediction.

This is an ambitious research program that raises a number of difficult challenges… It requires trained man power and time, and therefore extensive resources that **Hi! PARIS** is helping to mobilize.

AI for Computer Graphics an example of co-evolution

Artificial intelligence, and more specifically machine learning, has recently spread to Computer Graphics (CG). In specific cases, Deep Learning has improved inverse modeling and simulation—i.e., combining laws on shape or motion with user control, a long-lasting problem in this field. While purely generative neural networks—e.g., **Generative Adversarial Networks** (GAN), trained to generate many examples similar to those in a learned dataset—are not directly usable in CG due to their lack of control, they have been adapted to solve specific inverse procedural modeling problems such as generating a realistic terrain from sketched ridges and rivers.

Deep Learning solutions have been proposed for modeling 3D faces from a sketch and adapted to learn and control fluid simulation. Deep Learning has also been massively used, with Reinforcement Learning, for character motion control. A combination of machine learning techniques has enabled the robust design of chain reaction layouts, and the capture of Laban’s motion qualities expressing the “mood” of a motion from interactive hand gestures, then transferred to an animated character.



Designing a "creative testbed" for scientists
enabling them to refine and interact with their visions

Could digital images be turned into a creative media, more expressive but as simple to use as a pen, enabling scientists to directly express these visions in the form of interactive 3D environments, manipulate them, test different hypotheses through interaction and inject new data on the fly, in order to get inspiration and refine their thoughts? This is the goal that Marie-Paule Cani, professor at **Hi! PARIS**, is pursuing. If successful, it will make scientific thinking directly possible in digital form, from early intuitions to progressive refinement, testing, and finalization of an idea.

Recent advances in Artificial Intelligence are paving the way to just this kind of revolution in content creation. However, AI should be revisited from a **user-centered perspective**. More precisely, we aim at drawing inspiration from traditional art, where creation is typically achieved from coarse to fine scales. Starting from a white board, a user—for instance, a scientist with a specific phenomenon in mind—should be able to **sketch** a few elements, put them in motion in a few gestures, add constraints (e.g., constant volume deformations), explore the model in **3D**, refine some of the shapes and, in a click, transfer similar details to the other ones, refine motion as well, or add other constraints, freely zoom in or out to refine the model at a different scale, etc. He/she may also want to use analogies, through a "pipette" tool to learn and import either the distribution of details, their visual aspect, a velocity field, or a deformation field, from some real data, possibly of a different dimension, such as a video. Scripting motion to enable narration, through the specification of different stages for the phenomenon, should also be enabled.

At each stage, this user should be able to stop the **modeling** process and interact with the just-created virtual testbed, for instance through the application of

> *Visual representations for science do not fully exploit the current capabilities of computers*

user-controlled interaction forces, and possibly "feel" the reaction of the environment through a force-feedback device. He/she should be free to apply any of these design, exploration, and interaction operations in arbitrary order.

Marie-Paule Cani is a Computer Science Professor at École Polytechnique since 2017, and President of the CS Department since 2020. She chairs the Scientific Council of the French Computer-Science Society (SIF). Her research in Computer Graphics has focused on Shape Design, Computer Animation, and Smart Models enabling the expressive design of animated virtual worlds.

Involved in the committees and editorial boards of the main conferences and journals in her field, she has received several awards for her research including the Irène Joliot-Curie Mentorship award (2007), a silver medal from CNRS (2012), and the Eurographics Distinguished Career Award in 2022. She has been elected to the Academia Europaea (2013), the ACM SIGGRAPH Academy (2019), and the French Academy of Sciences (2019).

She also has been working for around a decade with TotalEnergies (to discover mountain forms or to reveal water sources) and L'Oréal (providing new scales to explore hair or skin) to transform her theories and research into practical tools.

• **Hi! PARIS Fellowship**

CREating and Animating Testbeds of Ideas, through Virtual Environments with Artificial Intelligence

This **Creative AI** ambition is a highly unconventional project, compared with mainstream research in AI and Graphics. In contrast with intelligent systems that aim at tackling creation for us, the goal is to support and enhance users' own creativity. As shown by the state of the art, none of the existing approaches provides an interactive, ready-to-use media to support creative, scientific thinking.

To achieve this, Graphic Design and Artificial Intelligence methods need to be revisited from a user-centered perspective, as promoted by modern Human-Machine Interfaces. This will imply bridging the gap between (i) Expressive gesture-based creation, based on sketching, sculpting or smart copy-paste, developed for isolated 3D shapes so far; (ii) Procedural models for visual simulation, conveying prior knowledge through rules and constraints, but often slow and difficult to control; and (iii) Learning processes that build on data and examples to quickly calibrate models or generate contents though analogies.

Put the user in the loop

In terms of the AI technologies used, while we will take inspiration from inverse simulation and perceptual transfer methods based on deep learning, we plan to focus on lighter and more versatile learning solutions, designed to put the user in the loop during the creation phases and allowing them to learn from sparser datasets.

Benefits for science and society

We are seeking the development of an AI that could make us, humans, more creative. The expectation is to provide scientists with an interactive, 3D modeling testbed that may serve as support for thinking about, toying with, and exploring novel ideas. Since the system will automatically handle tasks such as the replication of details or ensuring that constraints on shapes or motion are met, we could see this as a system enabling co-creation tasks, where the human is the leader and AI the helper, as recently explored in musical creation.

This work should bring important benefits to science, technology & society, by:

- offering the ability to design and interact with consistent, 3D environments to ease scientific thinking and accelerate research through easier communication and collaboration between researchers;
- in the CG domain, cracking the long-standing problem of modeling each natural phenomenon one by one, with specific models that cannot often be reused;
- bridging the gap between real images and user-created examples, by extending learning to sparse and diverse datasets;
- enabling the use of prior knowledge and machine learning techniques to support each other for solving a complex problem, rather than competing.

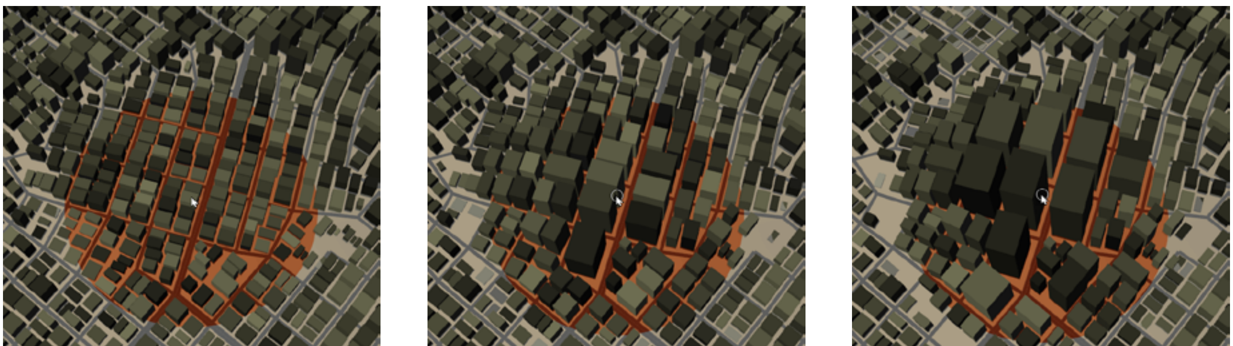
The resulting virtual testbeds could be used by scientists, but also be adapted to engineering needs, and provide, in a simpler form, an educative playground for the public willing to experiment and create. In summary, it will provide not only hands-on experimentation, but also means of communication and education, impacting at the same time the nature of technologies and their usage.

References

X. Zhou, P. Chang, M.-P. Cani, and B. Benes, *"Urban Brush: Intuitive and Controllable Urban Layout Editing,"* in UIST'21: The 34th Annual ACM Symposium on User Interface Software and Technology, Virtual Event, United States, Oct. 2021, vol. 40, no. 4, pp. 796–814.

P. Ecormier-Nocca, G. Cordonnier, P. Carrez, A. Moigne, P. Memari, B. Benes, M.-P. Cani, *"Authoring Consistent Landscapes with Flora and Fauna,"* ACM Transactions on Graphics, Aug. 2021.

See figure below: applying the UrbanBrush "Attractor", with Merge+Split atomic operation.



Decision-making in the age of AI

Algorithmic decision-makers have both economic and societal impacts

An algorithm is a series of operations commonly executed by computers to automatically perform data processing, predictions and reasoning. Given their abilities to process "big data" and perform these tasks consistently, algorithms are expected to outperform humans in many circumstances. We therefore see vast implementations of **algorithmic recommenders**, **advisors**, and **decision-makers** in business and society.

Examples of algorithmic recommenders include recommendation systems that produce algorithm-based product recommendations on Amazon and movie recommendations on Netflix. This category of algorithm-based applications is relatively mature and widely used in business practice. Examples of algorithmic advisors include automatic systems that produce algorithmic diagnoses of medical symptoms and signs, and algorithm-based judgmental opinions to assist humans in their judgment.

Algorithmic recommenders, advisors, and decision-makers represent the three broad categories of algorithm-based applications and systems that are used in practice

While the first two categories are used to assist humans in their judgment-making and decision-making, the more recent advances in artificial intelligence and machine learning technologies aim to enable automated algorithm-based decision-making on behalf of humans (so called algorithmic decision-makers), with vivid examples of algorithmic trading systems which execute orders using automatic pre-programmed trading stock instructions and the recent examples of autonomous vehicles which drive without human intervention. Algorithmic recommenders, advisors, and decision-makers represent the three broad categories of algorithm-based applications and systems that are used in practice.

Xitong Li, associate professor at HEC Paris, examines the impact of these three categories of algorithm-based applications on our business and society. The following questions arise:

- *How does algorithmic investment influence human investment?* Despite the considerable body of literature on social influence from humans to humans, no research has investigated the extent to which humans' decisions reflect the influence of prior decisions made by algorithms, referred to as "algorithmic influence". We will examine these algorithmic decision-makers specifically through algorithmic investment robots in an online peer-to-peer lending market.
- *Does presenting prediction performance matter? How does algorithmic transparency influence adoption of algorithmic advisors?* A recent stream of experimental studies finds that individuals have a tendency toward algorithm appreciation, in the sense that they exhibit a higher degree of advice taking when the advice comes from an algorithmic advisor than from humans, even when the advice is identical. Yet, the research also shows that the extent of algorithm appreciation depends on contingent factors, such as an indication of the algorithm's prediction performance.
- *When is it valuable to present retargeted recommendations vs generic product recommendations? How do recommender systems lead to consumer purchases? How beneficial are recommendations to consumers?* This stream of questions reflects three gaps in the literature on algorithmic product recommendations, and aims at examining algorithmic recommenders, and more specifically algorithmic product recommenders in online retailing.

▲ *The transparency of an algorithm refers to the degree to which the information related to the algorithm is revealed and also how such information is presented. The information may be about the algorithms' prediction performance or about the input variables used by the algorithms.*

Through these topics, we aim at **Hi! PARIS** to build a viable research community that is interested in exploring the economic and societal impacts of algorithm-based applications.

References

X. Li, J. Grahl, and O. Hinz, *"How Do Recommender Systems Lead to Consumer Purchases? A Causal Mediation Analysis of a Field Experiment,"* Information Systems Research, doi: 10.1287/isre.2021.1074.

S. You, C. Yang, and X. Li, *"Algorithmic Versus Human Advice: Does Presenting Prediction Performance Matter for Algorithm Appreciation?"* Forthcoming in Journal of Management Information Systems.

• **Hi! PARIS Fellowship Starting**

The Impacts of Algorithmic Decision-makers, Advisors, and Recommenders on Business and Society

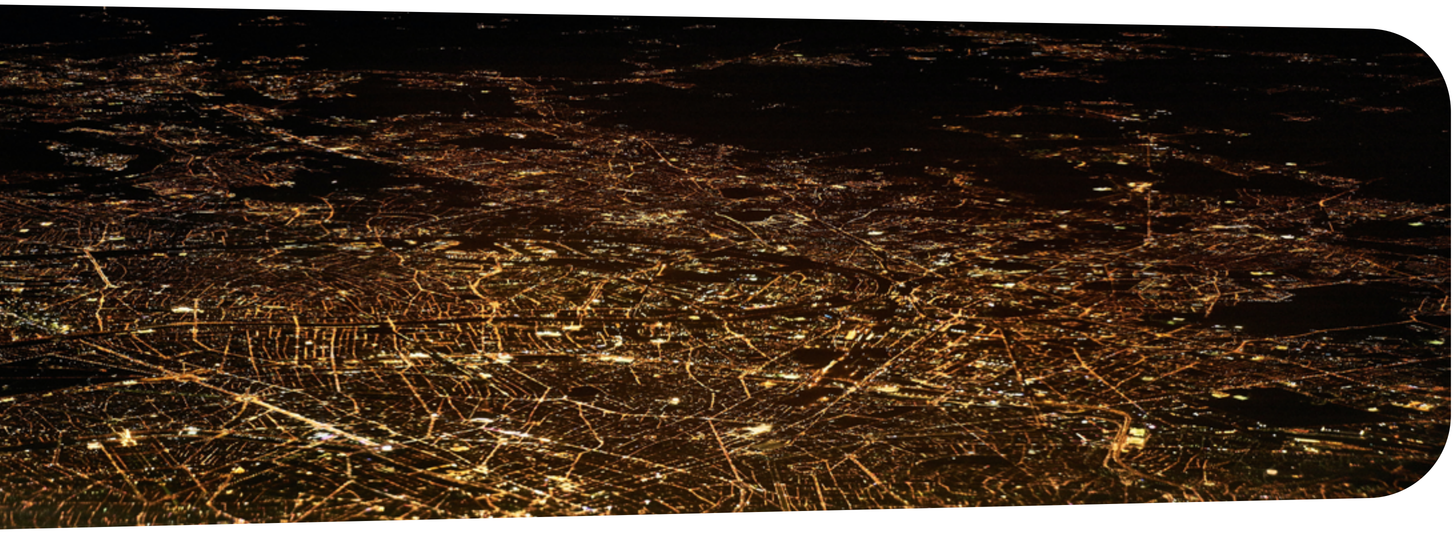
Xitong Li is an Associate Professor of Information Systems at HEC Paris, France. His primary research interests are the economics of information technologies, including social media, crowdfunding, digital marketing, online education, and AI. His primary research methods include applied econometric analysis, field and laboratory experiments. He works with collaborators from both academia and industry, across various countries (France/Europe, the U.S. and China).

Xitong's research appears in leading international journals, such as Information Systems Research, Management Information Systems Quarterly, Journal of Management Information Systems, and various ACM/IEEE Transactions. Xitong's research has been granted by ANR AAPG France (solo PI), equivalent to National Science Foundation (NSF) in the U.S. for 2018-2023. Xitong currently serves as an Associate Editor for Information Systems Research.

Addressing sequential decision-making problems related to energy transition
using reinforcement learning

Reinforcement Learning (RL) is a type of learning process where an intelligent agent (i) is interacting in a sequential manner with an unknown environment, (ii) aims to maximize its cumulative rewards and (iii) uses function approximators to generalize the information acquired from the agent's interaction with the environment. These techniques have greatly benefited from deep neural networks and have achieved multiple successes when combined with such function approximators.

Energy transition is a field where sequential decision-making problems are complex and sometimes beyond the capacity of existing RL techniques. This is the area in which Professor Damien Ernst conducts his current research work, either developing new RL algorithms able to tackle this challenge, or adapting existing ones.



• **Hi! PARIS Chair holder**

Reinforcement learning applied to Energy

Professor Damien Ernst is Invited Professeur at Télécom Paris.

He is also a Professor at the Department of Computer Science and Electrical Engineering at the University of Liège and the Chief Scientific Officer at Haulogy, a company specialized in the creation of intelligent computer software for the energy sector. Professor Ernst is well known for his work in reinforcement learning, a subfield of artificial intelligence, and the application of these techniques to the many decision-making problems faced by the electrical industry.

In recent years he has worked on the computation of long-term energy planning problems and has advocated using computational techniques for the creation of several new disruptive models (an electrical global grid, remote renewable energy hubs, etc.) for speeding up energy transition. He has received numerous awards for his work, including the prestigious Blondel Medal in 2018.

https:// www.damien-ernst.be/

Studying and designing new reinforcement learning algorithms to interact in electricity markets

Electricity markets are complex systems which have been specifically designed for trading electricity. These markets are continuously being adapted and optimized to better fit with the multiple ways electricity, and in particular renewable energy, is generated, by providing greater possibilities to trade electricity closer to real-time generation. This is particularly important in the context where more and more electricity is generated by renewable sources, but whose production level cannot always be accurately predicted, even twenty-four hours in advance.

Interaction with these markets can be formalized as **highly stochastic decision-making problems with high-dimensional action spaces**. RL techniques are well-suited to designing trading algorithms which are able to interact with them, especially when traders value flexibility (i.e., they have, for example, batteries in their portfolio) since flexibility usually implies time-coupling constraints between their different trading decisions. There is also an increasing willingness from consumers, even those at the residential level, to buy or sell their—usually photo-voltaic panel-generated—electricity in electricity markets without going through a retailer. Very few of these "minor players" have a high-level of trading expertise and in such a context, reinforcement learning algorithms could be used to design small intelligent trading agents which would effectively help them to trade their demand for or generation of electricity efficiently.

Combining computation of optimal environments and policies for reinforcement-learning problems

Most of the work involving RL focuses on the computation of **policies** which are able to take action that will lead to a high reward in a given environment. While many engineers are interested in such policies, they often face the combined problem of computation of the right environment and of the right policy.

Let us consider, for example, an engineer who builds a robot, say, to do the housework. He has of course to create the right policy for controlling the robot (i.e., for controlling its arms, its legs, etc., in the case of a humanoid robot) and also needs to make design choices for the hardware parts of the robot (i.e., how long the arms or legs should be). These choices will define (part of) the environment the intelligent policy will have to control. Hence there is a need to optimize a combination of the environment and the policy to achieve optimum performance and results.

In the field of RL, such **optimization problems** are poorly studied. We note that they are also of great importance to distributed energy systems (micro-grids, renewable energy communities, etc.) where one often needs to compute the best investments to be made, for example in batteries, photo-voltaic panels and/or small generators, while at the same time controlling those assets (i.e., when to charge/discharge the battery) in an optimal way.

Integration of renewable energy sources and batteries into electrical networks

with a special focus on distribution networks

Nowadays, electricity networks and, in particular distribution networks, face multiple challenges created by windmills, photo-voltaic panels, and the new loads (i.e., heat pumps, electric vehicles, etc.) that are being connected to them. One well-known example of such problems is the voltage problem caused by photo-voltaic panels connected to a low-voltage network. Indeed, when the sun is shining and the load consumption is low, these panels inject a significant amount of power into the distribution network that is "pushed" to the medium voltage level through local voltage increases. These voltage increases threaten the security of the distribution network and need to be controlled.

This control is often accomplished by managing the power electronics of the photo-voltaic panels and acting on local load consumption by trying to shift consumption to the hours during which the panels are producing a lot of electricity. This is one among the tens, if not hundreds, of new control problems that network operators must address. They are often very complex problems due to aspects such as the size of the network, the difficulty of establishing good models, the poor observability of parts of the network, etc. Artificial intelligence techniques and, in particular, RL techniques can play a key role and are certainly generic tools that network operators are increasingly willing to exploit in response to the many challenges caused by the current energy transition environment.

Energy models to speed up

the energy transition

In parallel with this work in the field of decision-making, Damien Ernst develops a line of research related to the study of "out-of-the-box" energy models that could significantly speed up the energy transition. The tools used for these studies rely, among others, on classical optimization and artificial intelligence techniques. However, these studies are not limited to numerical results but also involve collaboration with engineers from various fields in the physical sciences (chemists, etc.), economists, and researchers in social sciences. Among these different models, two receive particular attention.

The first is the Global Grid, an electricity network spanning the entire planet and connecting most of the large power plants in the world. Such a network would allow for a natural smoothing of renewable energy fluctuations as well as the harvesting of renewable energy in places where it is abundant.

The second is the model of a remote renewable energy hub for Carbon-Neutral Fuel Production. These hubs are built around the idea of synthesizing renewable fuels where there is a lot of sun and/or wind, and transporting them back to demand centers. Recent advances in direct air carbon capture would even allow for the synthesis of carbon-neutral hydrocarbons in such places.

SEE ALSO

Damien Ernst and his collaborators developed recently a new modeling language and associated tool called the Graph-Based Optimization Modeling Language (GBOML). GBOML implements a hierarchical hypergraph abstraction of optimization problems. It provides language constructs to represent such hierarchical structure and facilitate the encoding of time-indexed models.

B. Miftari, M. Berger, H. Djelassi, and D. Ernst, *"GBOML: Graph-Based Optimization Modeling Language,"* Journal of Open Source Software, vol. 7, no. 72, p. 4158, 2022, doi: 10.21105/joss.04158.

Apprehending how individuals make decisions
when they use search engines

Roland Rathelot is an applied micro-econometrician at ENSAE Paris. His scientific approach is to combine large and novel datasets from administrative or online sources with state-of-the-art data science and econometric methods to answer questions about the way agents make search decisions. In joint research work conducted in collaboration with *Arbetsförmedlingen*, the Swedish Public Employment Service, he uses data on clicks on Platsbanken—the biggest online job board in Sweden—to design a recommendation system based on collaborative filtering with implicit feedback.

Designing and evaluating recommendation systems for job seekers

Applying standard algorithms to improve labor-market outcomes is not straightforward: These algorithms tend to recommend the vacancies that are most popular. Because jobs cannot be shared across workers, it is necessary to choose hyper-parameters that balance the relevance of the recommendations with the coverage over the pool of vacancies, by giving bigger weights to vacancies that are relatively newer, and have not yet received many applications.

In order to evaluate the impact of the recommendation system on the labor market, the research team used a two-sided randomized controlled trial applied to real-time data on vacancy postings and job ad views during the COVID-19 pandemic. Half of the users were assigned to a control group for which recommendations were not available, when at the same time, half of the vacancies were randomly excluded from recommendations. This design allowed an evaluation of the impact of the recommendation system from the point of view of both job seekers (job search activity, labor-market outcomes), and employers (e.g., number of applications received, probability of re-posting a vacancy).

The recommendations provided by the algorithm—for each job seeker, a set of 10 recommended job postings on which they had not clicked yet—were available to all eligible Swedish job seekers between March 2020 and April 2022.

Reference

L. Hensvik, T. L. Barbanchon, and R. Rathelot, *"Job search during the COVID-19 crisis,"* Journal of Public Economics, vol. 194, p. 104349, 2021.

• Hi! PARIS Chair holder

Leveraging data science to improve matching on labor markets

Roland Rathelot is Associate Professor at ENSAE Paris. The most important application in his current research agenda is about search on labor markets, for which he collected detailed new data on individual search behavior. Some of his work is also dedicated to producing new empirical techniques that allow applied researchers to overcome methodological issues posed by the structure of the data that are becoming available.

He is a research affiliate at the Jameel Poverty Action Lab (J-PAL) and the Centre for Economic Policy Research (CEPR), an external research fellow at the Centre for Research and Analysis of Migration (CReAM) and a research associate at the Centre for Competitive Advantage in the Global Economy (CAGE, Warwick).

http://rolandrathelot.com/

Matthew Yeaton is an Assistant Professor at HEC Paris, focused on topics at the nexus of organizational culture and social network structure. As a strategy scholar and computational social scientist, Matthew uses AI tools such as Natural Language Processing (NLP) to analyze language and culture in social networks in order to understand how cultures and knowledge are transferred. For example, he investigates how network interventions can disrupt hate speech online without resorting to banning users.

He obtained his PhD in Management from Columbia University. Prior to starting his PhD, he worked as a senior research analyst at the Federal Reserve Bank of New York.

See also page 34.

- **Hi! PARIS Chair holder**

Strategy and Computer Science

Studying how the information environment affects strategic decision-making

Organizations and individuals can affect the information environment in a range of ways, including **persuasion**, **positioning**, and **information manipulation**. Any setting where the information environment can be manipulated provides an opening for strategic action.

The interplay between strategic choice and the information environment, including how strategies morph as the environment changes, is a dynamic that Matthew Yeaton, assistant professor at HEC Paris, is currently studying in early-stage work through two different use cases.

Fake news via intentional distraction

Presidential elections are an ideal experimental time for people who wish to frame public opinion. Numerous methods have been used and tested in several countries, and remain active, to influence electors via social media, especially Twitter.

> *The strategic influence of information can be valuable for organizations and individuals alike*

In this context, and with the Mexican 2010 presidential election as a backdrop, we developed a formal model of fake news as a **bias** of the relative precision of signals rather than a bias of the signal means. Under this type of bias, Bayesian synthesis of signals can be biased (compared to the true model) even though none of the individual signal means is biased.

This model can help explain a strategy of diffusing “white noise” on Twitter. A small army of Twitter trolls manage the image of a candidate on social media. They do not use actively hostile messages for their opponents, neither do they promote their candidate directly. Instead, they use a white noise strategy that magnifies the noise during times of particularly bad news about their candidate. This type of fake news is realistically possible only in the wild west of digital information diffusion.

Discerning the desirability of a product

While there are some industries where the marketplace may have perfect information about product desirability, for others the marketplace may have at best a hazy or biased belief about desirability. Firms differ in status and in customer beliefs about desirability, and differences in status or beliefs may also drive positioning choices if these positioning choices have the power to affect beliefs.

We model this problem as an incomplete **information game**. Our formal results reveal how information quality and costs impact firm positioning choices and industry-level positioning heterogeneity, and how these change over time in reaction to changes in consumer access to information about desirability.

In ongoing work, we test our predictions by leveraging shocks to the cost of information stemming from the staggered UK broadband roll-out.

Reacting promptly
to change

Developing machine learning and artificial intelligence methods, and exploiting alternative sources of information, to better forecast macroeconomic and financial activity in real time, is crucial for society since it allows actors to react promptly to changes in the economy with appropriate counter-cyclical economic policies.

The nowcasting quest

At **Hi! PARIS**, we develop machine learning tools tailored for **a time-series environment** with the aim of nowcasting macroeconomic and financial aggregates by using alternative data—for instance Google Search data—together with official data.

In a multidisciplinary project that involves tools from Machine Learning, Artificial Intelligence, Econometrics and Economics, Anna Simoni, professor at ENSAE and École Polytechnique, has been pursuing a line of work on these topics for several years.

Nowcasting aims at providing an evaluation of **macroeconomic** and **financial** activity in real time, which is crucial for the policymaker to react promptly to changes in the economy with appropriate counter-cyclical economic policies. Indeed, the problem that forecasters face is that official series are often published with a delay. We are interested in exploiting in the best possible way the economic informational content of Google search data and in analyzing how the combination of these data with the official series (when available) can improve prediction accuracy. Because Google Search data contain an ultra-high number of variables (categories) compared with the time dimension, our first goal is to shape pretesting procedures that retain only the more important Google search variables to improve nowcasting accuracy. This work builds on several previous contributions and is carried out with new sets of alternative data. We focused first on the use of **Google search data** to nowcast euro area Gross Domestic Product (GDP) growth, develop-

• **Hi! PARIS Fellowship Consolidator**

Macroeconomic Nowcasting With High-Dimensional Google Search Data: Theory and Practice

Anna Simoni is Senior Researcher at CNRS / CREST and Professor of Econometrics and Statistics at ENSAE and École Polytechnique. Before joining the CNRS, she was Assistant Professor of Statistics at Bocconi University, Milan, between 2009 and 2012. She has been associate researcher at the University of Mannheim. She also held the post of Visiting Assistant Professor at Boston College.

Bronze medal CNRS (2019), Visiting Assistant Professor at Boston College (2014), Fernand Braudel Fellow by the Department of Economics at the European University Institute in Spring 2020, Associate Editor of Journal of Econometrics (2021-), Fellow of Institut Louis Bachelier (2021-), PhD at Toulouse School of Economics.

ing an estimator that combines variable preselection and Ridge regularization and studying its theoretical properties. More generally, we also lead theoretical and empirical studies of various machine learning techniques coupled with a pretesting step to build nowcasts by exploiting alternative sources of information.

The aim is to establish the out-of-sample theoretical properties of procedures based on different penalizations and factor-based methods to nowcast macroeconomic aggregates. These properties will take into account the nature of alternative data, their temporal dependency structure and correlation in the data.

A further objective is to show that Google search data can be useful to detect structural breaks, and to propose a procedure to estimate the break points and the coefficients of the corresponding models. A final goal consists in reducing the mean squared forecasting error of the nowcasts through bagging methods that are developed for this particular framework involving nowcasting and time-series environments.

Nowcasting, providing real-time evaluation of macroeconomic and financial activity, is crucial for policymakers

Ongoing work

To date, we have established theoretical results concerning the targeted preselection in a framework of ultra-high dimensional data, and the data-driven choice of tuning parameter. From an empirical point of view, we have used Google search data to nowcast GDP for the euro area, Germany and the U.S., and have shown that Google search data allow us to improve nowcasting accuracy if appropriately preselected.

This work is extended to compare information coming from two sources of alternative data in terms of nowcasting accuracy to nowcast U.S. GDP. The alternative data considered are Google search data and the news attention monthly series obtained from The Wall Street Journal.

References

L. Ferrara and A. Simoni, *"When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage,"* arXiv, 2020. doi: 10.48550/arxiv.2007.00273.

M. Mogliani and A. Simoni, *"Bayesian MIDAS penalized regressions: Estimation, selection, and prediction,"* Journal of Econometrics, vol. 222, no. 1, Part C, pp. 833–860, 2021.

The Ridge after Model Selection estimator, as defined in "When are Google data useful to nowcast GDP? An approach via pre-selection and shrinkage", where $\alpha > 0$ is a regularization parameter that tunes the amount of shrinkage and $\widehat{M}_g$ is the set of preselected variables. ▼

$$\widehat{\beta}^{(w)}(\alpha) := \underset{\beta\in\mathbb{R}^N;\beta_{g,j}=0,j\in\widehat{M}_g^c}{\operatorname{argmin}} \left\{ \frac{1}{T}\sum_{t=1}^{T}\left(Y_t-\beta_0-\beta_s' x_{t,s}-\beta_h' x_{t,h}-\beta_g' x_{t,g}\right)^2+\alpha\|\beta\|_2^2 \right\}$$

Practicing inclusiveness through innovation law

Law is evolving under the pressure of technology and innovation. In turn, it thus has the potential to produce change and generate a socially and economically inclusive digital economy, for both business and society at large.

In the field of innovation law, one topic of interest at the frontier between law and society is to examine how legal institutions and technologies can support practices of innovation in a socially and economically inclusive way. What is at stake is how to disseminate to firms the opportunities for producing at the frontiers of innovation and ensuring that the benefits of advanced technologies can be extended to society at large in a lawful and ethical way.

This is the path that Pablo Baquero, assistant professor at HEC Paris, explores. He develops an interdisciplinary approach at the intersection between law and innovation, which entails an understanding of the internal operation of technologies (e.g., algorithms, blockchain, smart contracts), as well as the ability to design them for legal compliance. By providing an understanding of the inner workings of technical applications, his research helps to give a new meaning to the role of law, unhinged by its disciplinary frontiers.

Computational methods to assess lawyers' performances

In a collaboration involving an interdisciplinary group comprising legal scholars and computer scientists within **Hi! PARIS**, we examined the potential of artificial intelligence methods to generate relevant information about lawyers' performances in the courtroom. These methods have the potential to generate relevant information for consumers—arguably, to minimize their information **asymmetry** in relation to providers of legal services—and to provide law firms with reliable methods to assess legal professionals' performances.

Such computational methods are increasingly being used in practice, in parallel to traditional performance indicators for the legal profession (i.e., client-based and peer-to-peer based indicators). They are believed to overcome some of the flaws presented by these traditional methods (such as the use of anecdotal evidence, the existence of a selection **bias**, and the lack of **transparency**). Currently, however, their use presents two important shortcomings.

> *Computational methods still have important shortcomings*

First, computational methods to assess legal professionals' performances are available mostly through private companies that charge a significant fee to those interested in having access to this information. As a result, most consumers and small and medium firms do not have access to information about legal professionals' performances generated through computational methods.

Second, the software used by these companies and their underlying source code remain a "black-box"—there is very little or no information provided by private companies about the inner workings of these technologies and how they generate predictions/recommendations.

With the purpose of overcoming these limitations, we conducted a case study to build legal analytics based on the legal decisions of the French Courts of Appeal. We drew on a dataset of 40,000 French court decisions derived from the Légifrance database and further analyzed a subset of 8,045 cases from the Court of Appeal through methods such as **machine learning, text mining, and natural language processing**.

More specifically, our results demonstrate how and what kind of information the computational indicators supported by AI methods can generate about the experience of lawyers in litigation, combining information about the number of win-losses, the particular field of law where that experience has unfolded, the types of collaborators and opponents that have been faced in the course of litigation, and an indicator regarding the difficulty of the case.

We studied the resulting analytics, seeking to highlight their potential and limitations in building indicators regarding the performances of legal professionals in the particular context of French law.

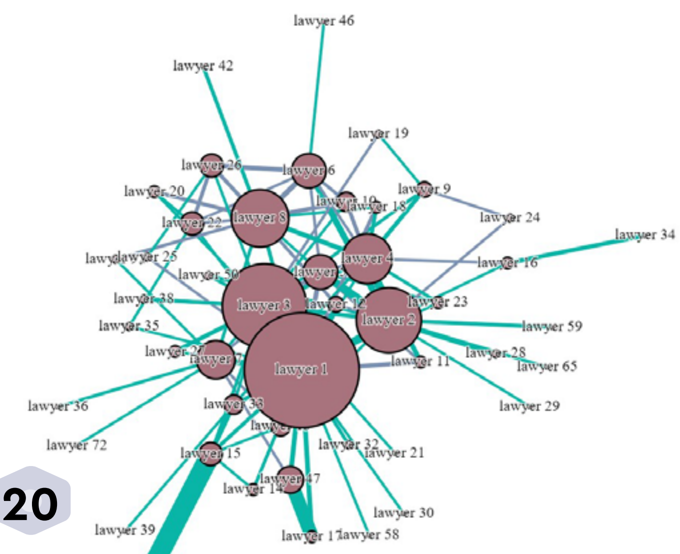
Through this case study, we shed light on the methodological steps followed in building computational indicators, evidencing the challenges involved in that process and seeking to break open the black-box that envelops the production of these legal analytics.

We also proposed several adjustments to computational analytics to address methodological aspects, such as developing more sophisticated win-loss ratio rankings to evaluate legal performance in the courtroom; combining different types of information; expanding the methods and criteria of evaluation of legal services to measure the competence of different types of lawyers; and integrating representatives of the legal profession in the development and/or revision of these types of algorithms.

REFERENCE

D. Restrepo-Amariles, P. Baquero, P. Boniol, R. El Hamdani, and M. Vazirgiannis, *"Computational Indicators in the Legal Profession: Can Artificial Intelligence Measure Lawyers' Performance?"* (June 30, 2021). Journal of Law, Technology and Policy, Vol. 2021, No. 2, 2021.

Below : Network of lawyers that have collaborated in trials.



• **Hi! PARIS Chair holder**

Law and Smart Contracts

Pablo Baquero is Assistant Professor at HEC Paris. He conducts research with a group of legal scholars and data scientists on artificial intelligence and legal issues concerning ethical innovation. He recently published his first book, "*Networks of Collaborative Contracts for Innovation*", examining how companies willing to pursue collaborative innovation projects structure their contractual agreements. He also investigates how the "black box" of algorithms could be designed to comply with law and regulations.

Pablo holds a PhD in Law from the University of Cambridge, a LL.M. from Harvard Law School and a LL.B. from the Federal University of Rio Grande do Sul, Brazil. He used to work as a lawyer in France and in Brazil, dealing with commercial arbitration, contractual transactions, and foreign investments. He has also been a consultant for the Doing Business Project by the World Bank of Washington DC.

P. Baquero, *"Networks of Collaborative Contracts for Innovation"*, International Studies in the Theory of Private Law, Bloomsbury Publishing, 2020. ISBN: 9781509929986

Creating AI with rules

Since 2020, there has been a consensus on a series of AI principles around eight main themes: privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, and promotion of human values. These principles are at the core of guidelines such as the Recommendation on the Ethics of Artificial Intelligence adopted by UNESCO's General Conference at its 41st session, or the EU Artificial Intelligence Act, which is still under discussion.

With its world-class researchers from the technical, managerial, legal, economic, ethical and societal domains, collaborating with teams all around the world, **Hi! PARIS** is equipped to invent the AI solutions which comply with present and future rules, here and elsewhere.

Towards trustworthy AI

AI trustworthiness has been recognized as a major prerequisite for people and societies to use and accept such systems. In April 2019, the High-Level Expert Group on AI of the European Commission defined the three main aspects of trustworthy AI: it should be lawful, ethical and robust. Providing a warranty on this topic is currently a matter of study and discussion.

To prevent deep models from capturing biased features and operating on biased datasets, Enzo Tartaglione devised a specific regularization strategy (see page 26). New models of learning, such as federated learning (studied by Aymeric Dieuleveut, see page 36) which distribute data on decentralized storage, do their bits for privacy preservation.

Privacy is a concern for both individuals (concerned with the protection of their personal data) and firms (facing legal constraints preventing them from exploiting valuable data)

Complying with regulations

How can controllers, processors and subprocessors of personal data ensure that their processing of the data that flows through these different companies comply with data protection regulations?

This is what Pablo Baquero and his collaborators call "*compliance in the data supply chain*" (see next page), and all stakeholders in this chain must be aware of the need to elaborate a Data Protection Impact Assessment in their activity.

Need for further knowledge

Policymakers worldwide are increasingly restricting (online) user tracking to protect user privacy. Their activities include, for instance, the General Data Protection Regulation (GDPR) in the European Union, the California Consumer Privacy Act (CCPA) in the United States, and China's Personal Information Protection Law (PIPL). Web browsers, such as Firefox, accompany these activities by also restricting user tracking.

All stakeholders must re-evaluate their business model, but the knowledge required to make the right choices is still lacking, as Klaus Miller (see page 23) emphasizes.

SEE ALSO: LAW, SOCIETY & AI SEMINAR

The "Law, Society & AI" seminar is a series of invitational talks by researchers on questions at the intersection of legal, societal, and artificial intelligence issues.

Organized by David Restrepo Amariles (HEC), Winston Maxwell, Michalis Vazirgiannis, and Fabian Suchanek (all Institut Polytechnique de Paris), it takes place online, as well as the premises of **Hi! PARIS** members. https://suchanek.name/work/research/lsai/

The seminar is supported by the NoRDF chair at Télécom Paris, a scientific project that aims to model and extract complex information from natural language text, and seeks to enrich knowledge bases with events, causation, conditions, precedence, stories, negation, and beliefs.

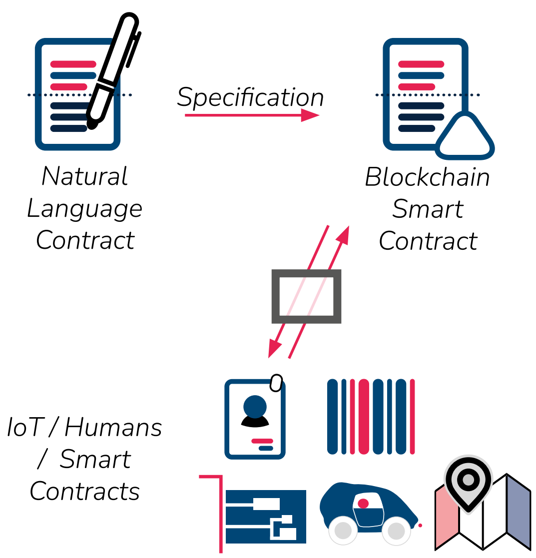
Privacy
a social demand and a market differentiator

Using smart contracts to ensure compliance with data protection laws in contractual supply chains

Most research on compliance with privacy laws is focused on the data subject: the rights and possibilities of enforcement that data subjects have against controllers for breaches concerning their **personal data**. It often involves research examining the privacy policies of companies, and their clarity and transparency for consumers, for instance.

In research conducted with computer scientists from the University of Ottawa, Pablo Baquero (see page 20) and his collaborators shift this perspective. Instead of focusing on the data subject, the spotlight is turned on how controllers, processors and subprocessors of personal data can ensure that their processing of the data flowing through different companies can comply with data protection regulations. This is what is known as "compliance in the data supply chain". The argument is that individuals' right to privacy will not be effective unless it is possible to monitor and hold processors accountable for data flows in the data supply chain. Nowadays, even if companies controlling and processing data want to comply with data protection laws and have a good privacy policy, it might be challenging to achieve this objective because of the complexities of monitoring the data flows between different companies.

In the context of this work, experimental research is being conducted on a case study related to the data supply chain of a network of leading IT companies in Europe.



Pablo Baquero's research in this book is conducted in the Smart Law Hub at HEC Paris, a research group that unites professors and researchers across institutions and disciplines to advance knowledge on the technological transformation of law. Their vision is that a new Scientific, Mathematical, Algorithmic, Risk and Technology driven law (SMART-Law) is emerging behind the back of black letter law, codification, and courts.

The driving forces propelling these changes are advances in social sciences, statistics, big data and machine learning. The Smart Law Hub believes that, for legal professionals to remain central players in an increasingly code-bound and risk-based society, it is necessary to get to grips with the challenges posed by SMART law from multiple perspectives by combining managerial, legal and computer science approaches.

Estimating the economic value of user tracking

Online publishers face difficulties evaluating the sustainability of their ad-supported business model that finances their content. Often this content is free of charge, which is particularly attractive for lower-income users. Advertisers need to evaluate how strongly user tracking increases their targeting abilities. Finally, when deciding on **user tracking**, policymakers need to make a careful trade-off between firms' interest in making profits, e.g., creating jobs or valuable content, and users' interest in protecting their privacy, which often means a decrease in user tracking.

The **online advertising** industry largely relies on data from user tracking to improve online advertising by better targeting users and measuring ad performance. An inability to track users could make online advertising less attractive for advertisers and lead to lower prices that advertisers pay publishers for displaying their ads.

However, the size of the potential decrease is unclear, and speculations vary widely. For example, a group of representatives from the online advertising industry postulated a substantial decrease by outlining that the increasing unavailability of user tracking "*will bring the single biggest change to the advertising ecosystem [...]*" (IAB Europe 2022). Other prominent industry voices like Michael Zimbalist, former CMO of the New York Times, argue that the death of user tracking will not harm the industry. These opposing positions illustrate the heated discussion in the online advertising industry.

The findings of empirical and academic studies also differ strongly, with estimated decreases in ad impression prices of between 8% to 52%. Analytical studies even show that less user tracking could both increase and decrease ad prices. Thus, the effect of user tracking restrictions on the online advertising market is still unclear.

This lack of knowledge is unfortunate for the online advertising industry, users, and policymakers.

At **Hi! PARIS** we aim to provide such knowledge through an empirical study that covers a wide range of publishers, by answering the following questions and research objectives:

- *How much value does user tracking generate for publishers?* An inside look at the average effect across all publishers and all data categories;
- *How does the value of user tracking differ between publishers?* An analysis of the heterogeneity across publishers;
- *How does the value of user tracking differ between data categories?* An analysis of the heterogeneity across different data categories (e.g., a user's browsing history).

The motivation for doing so is the fact that knowledge about the different effects on publishers provides a better understanding of the discussion among publishers and provides guidance on their future strategies for adjusting to a world without user tracking.

On the other hand, knowledge about the different effects of data categories enables us to compare the intrusiveness of data for user **privacy** with the usefulness of data for the online advertising industry. Policymakers could, for example, restrict the tracking of data that is not important for publishers but that strongly decreases user privacy (and vice versa).

• Hi! PARIS Chair holder

The Economic Value of User-Tracking

Klaus Miller is Assistant Professor of Marketing in the Marketing Department at HEC Paris. His research interests are at the interface between empirical quantitative marketing, management economics, and information systems. Specifically, his research is concerned with pricing, advertising, and customer management issues in the digital economy. Methodically, his research is based on quantitative empirical modeling, applied econometrics, distributed statistical computing, causal machine learning, as well as large-scale field and lab experiments.

Published in top-tier academic journals (e.g., the Journal of Marketing Research, the Journal of Product Innovation Management, or the International Journal of Research in Marketing), as well as management-oriented journals (e.g., Marketing Review Sankt Gallen, GFK Marketing Intelligence Review), he often collaborates with industry to answer research questions at scale.

Constraints to overcome

> *Allowing practitioners to compute before starting an evaluation what they should expect in terms of the distribution of non-experimental bias*

Experimental vs. non-experimental approaches

There is a relative consensus among scientists about the superiority of experimental approaches in order to reach causal statements. However, for ethical, time, or budget reasons, it is not always possible to implement randomized controlled trials. Researchers and policymakers are often left with non-experimental methods to evaluate policies and interventions.

In an ambitious international project*, Roland Rathelot (see page 15) and his colleagues are gathering secondary data from numerous previous studies using randomized controlled trials with imperfect compliance (i.e., where not all individuals assigned to treatment were actually treated, or vice versa) to measure the bias of non-experimental methods and its distribution. Looking into the variance of the bias within a class of interventions and evaluations would show how uncertain the use of non-experimental methods is. They compare experimental to non-experimental estimates obtained using state-of-the-art econometric methods (e.g., based on lasso or random forests) which allow practitioners to remain agnostic about which covariates to introduce into the model and what the functional form between covariates and the outcome is.

This work has already aroused the interest of several founders and policymakers (e.g., the UK Department for International Development, Innovations for Poverty Action). It is a tool that will be sustained over time: a website will be set up and enriched as the number of studies included in the analysis increases. As time goes by, it will be possible to produce more precise estimates of the distribution of non-experimental bias in a given context.

The end goal is to allow practitioners to compute, before starting an evaluation, what they should expect in terms of the distribution of non-experimental bias. If the distribution is narrow enough, that means that there is relatively little uncertainty, and non-experimental estimates are likely to be valid enough to reach causality.

* partnerships are listed at the end of the document

New constraints on the learning process

Deep Neural Networks (DNNs) can solve extremely challenging tasks thanks to complex stacks of (convolutional) layers with thousands of neurons, in particular to solve computer vision-related tasks like image classification, object detection or image segmentation. Their success stems from their ability to learn from examples, not requiring any specific expertise and using very general learning strategies, based on loss' minimization.

In recent years, much interest has been devoted to two different aspects of deep learning and computer vision:

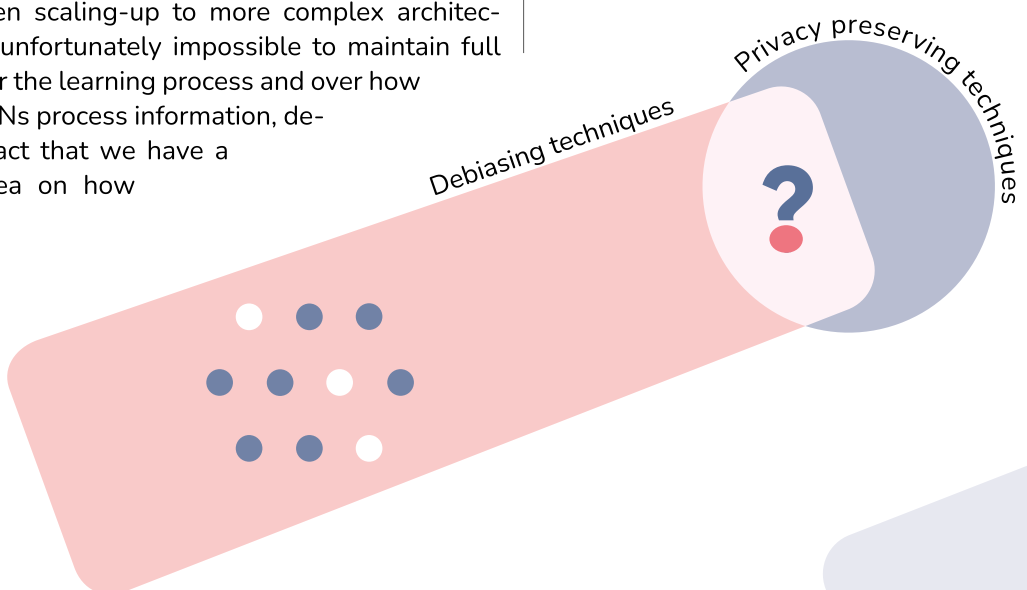
- *How can we improve the learning strategy of deep models? Have optimal features been extracted from these automatic models? Is there some unwanted information leakage? How can we make these models focus on extracting the right information from the input, in an automatic way?*
- *Do we really need all the complexity we are currently using to solve tasks with deep learning? Is there going to be any difference at training and at inference time? What about the environmental impact of AI training and deployment?*

In order to address the learning aspects of DNNs, many communities are hard at work, for instance pure information theorists and physicists. In particular, the latter community is able to describe learning properties on simple models (like the binary perceptron) using sophisticated mathematical tools, such as the Gardner analysis. However, these models intuitively have a limited learning capability to solve real-life tasks. When scaling-up to more complex architectures, it is unfortunately impossible to maintain full control over the learning process and over how exactly DNNs process information, despite the fact that we have a general idea on how they work. This is how DNNs ended up being named "black boxes".

However, there is the possibility of applying some lessons learned from the simple and applying them to the complex. We can design **regularization functions**, or in simple terms, we can add constraints to the learning process, in order to make DNNs satisfy constraints other than pure performance. We can enforce a proper feature selection in the learning process, or even power consumption minimization.

For example, designing a regularization term which employs an L0 metric on the parameters, and minimizing it besides the loss function, would allow us to both learn the target task and reduce the complexity of the model itself (resulting in power saving). The challenge here, however, would be to address a differentiable proxy as a good approximation for the measure to minimize, and to tune its weight properly during the learning process.

A fundamental research project is planning to define regularization strategies, to be coupled with the standard learning losses and rules, which will add constraints to the DNNs with respect to the features learned (and thus tackle problems like **debiasing** and **privacy preservation**, see page 26) and to the computational complexity required of the DNN model (and thus address **green AI** and **frugal AI** issues, see page 38).



Feature learning constraining

The so-called "unknown unknowns" in data are spurious relationships, which are mistakenly learned by a deep model. These make the model biased, meaning that it elaborates information in an unintended way while displaying high confidence in its predictions. This is a challenge in the quest for robust AI.

Such behavior has affected many recent works proposing AI-based solutions for COVID detection from radiographic images. Unfortunately, the available **datasets** at the beginning of the pandemic were heavily biased. This often resulted in models predicting COVID diagnosis with a high level of confidence, due to the presence of unwanted biases, for example by detecting the presence of catheters or medical devices for positive patients, their age (at the beginning of the pandemic, most ill patients were elderly), or even by recognizing the origin of the data itself (when negative cases were augmented by borrowing samples from other datasets).

Numerous works in the literature have attempted to address this issue. If the bias is known and can be explicitly manipulated, de-biasing directly at the data source is possible. In more challenging scenarios where this is not possible, adversarial and ensembling approaches are prominent, yet result in higher training complexity. The most challenging but effective scenario relies on training one simple model, without altering the data source but imposing extra constraints on the learning strategy.

A regularization strategy
to avoid biased features

Enzo Tartaglione proposes "EnD", a novel **regularization strategy** to prevent deep models from capturing biased features. In this framework, it is assumed that we know that the data might have some bias (like, in the case of COVID, the origin of data) but at the same time we are unaware of what it translates into (we do not have prior knowledge on whether the bias is the presence of a color, a specific feature in the image or anything else). EnD (standing for "Entangling and disentangling") regularizes the output of an "information bottleneck layer" within the deep model, where the regularizer:

- *entangles* the feature vectors extracted from data belonging to the same target class;
- *disentangles* the features extracted from data having the same "bias label".

Since the deep model is trained to minimize both the loss and EnD, all the biased features are discouraged from being extracted in favor of the unbiased ones. Compared to other de-biasing techniques, we have no training overhead: we do not train extra models to perform gradient inversion on the biased information or involve the use of generative adversarial networks, or even de-bias the input data. EnD works directly on the target model and is minimized via standard back-propagation.

Given that EnD entangles and disentangles feature vectors, it can be seen as a differentiable proxy for

Enzo Tartaglione holds a joint MS degree in Electronic Engineering from Politecnico di Torino and Politecnico di Milano, *cum laude*, in 2015. The same year, he also received a *magna cum laude* MS in Electrical and Computer Engineering at University of Illinois, Chicago. In 2019 he obtained a PhD in Physics at Politecnico di Torino, *cum laude*, with the thesis "From Statistical Physics to Algorithms in Deep Neural Systems". His main interests involve compression, sparsification, pruning and quantization of deep neural networks, privacy-aware learning, and data debiasing.

See also page 38.

• **Hi! PARIS Chair holder**

Regularizing deep neural networks: debiasing, efficiency & beyond

minimizing the mutual information between biased features and maximizing it between the same target features. In general, directly tackling the problem of mutual information minimization is hard, given both its non-differentiability and the computational complexity involved. Nonetheless, previous works have already shown that adding further constraints to the learning problem could be effective as, typically, the trained artificial neural networks models are over-sized and allow a large number of solutions for the same learning task. Our experiments show that EnD effectively favors the choice of unbiased features over biased ones at training time, yielding competitive generalization capabilities compared to models trained with other un-biasing techniques.

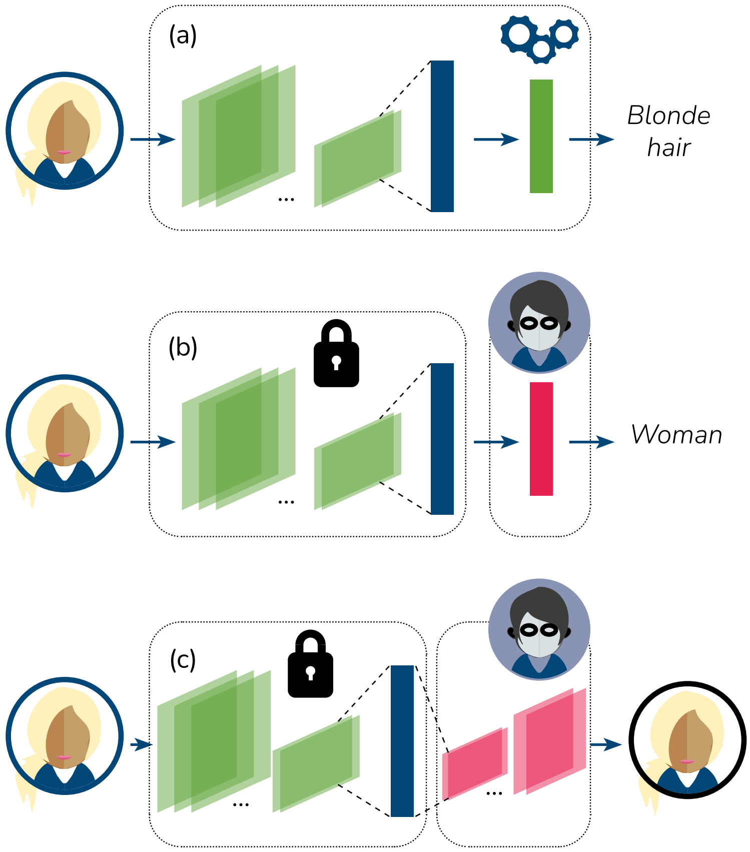
How can we be sure that sensitive information, given to a deep model, is not memorized or back-traced?

This technique paves the way to the exploration of scenarios where the **bias label** is an unknown: how to retrieve such information from trained, biased deep models? EnD certainly gives us a state-of-the-art tool to debias deep models (without modifying the input data), but the problem of detecting biases when they are not known remains open and worthy of exploration.

Data protection in mind

Another critical aspect to explore is the use of the disentangling term to prevent unwanted information from propagating. This is a known effect of deep models and of course, since the learning is governed by output error minimization only, there is no direct control over the model's decision process, not to mention the information extracted from the input and used thereafter.

Attacks can be conducted on a DNN model trained to solve some specific tasks, potentially revealing some private side information, not strictly necessary to solve the target task, yet leaking. This poses serious safety and data protection issues which are critical to the upcoming AI-centric metaverse, for instance. How can we be sure that our sensitive information, given to a deep model, is not memorized or back-traced? One objective is to design a regularization strategy able to hide, or even erase the sensitive information we do not wish to leak.



Standard training on some target feature, like hair color recognition (a), gender membership recovery (b) and input reconstruction (c). In red: The layer trained by the attacker to obtain the private information from the bottleneck layer.

FIGURE EXTRACTED FROM

C. A. Barbano, E. Tartaglione, and M. Grangetto, *"Bridging the gap between debiasing and privacy for deep learning."* Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.

REFERENCES

E. Tartaglione, C. A. Barbano, and M. Grangetto, *"EnD: Entangling and Disentangling Deep Representations for Bias Correction,"* in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Jun. 2021, pp. 13508–13517.

E. Tartaglione and M. Grangetto, *"A non-Discriminatory Approach to Ethical Deep Learning,"* in 2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), Los Alamitos, CA, USA, Jan. 2020, pp. 943–950.

Towards sustainable AI

Considering "*AI & the environmental stakes*" is a twofold problem: (i) we can try to tackle climate change with machine learning; (ii) we must work on more sustainable AI. Addressing climate change involves mitigation (reducing emissions), adaptation (preparing for unavoidable consequences), and tools for action (data visualization, social tools, etc.), all measures in which AI can help, in the knowledge that technology is not enough in itself. We must also be careful to apply AI in a way that does not make climate change worse.

At **Hi! PARIS**, we are ML practitioners and researchers committed to developing AI in an environmentally responsible manner.

Benefits of lighter models than Deep Learning

Rather than focusing on standard ML methods such as Deep Learning, the "smart 3D models"—seamless generation and editing of complex, animated 3D models, from a series of users' creative gestures (see page 8)—built by Marie-Paule Cani and collaborators, call for the development of lighter methodologies (green AI), enabling learning through a combination of prior knowledge and a small number of examples created on the fly by the user.

The combination of light learning and knowledge also generates explainable results, offering local user control.

Beware of pruning ANNs: this sums up to the total learning time.

It has become more and more difficult to size modern energy systems such as a Renewable Energy Community because of the complexity of the environment they are associated with (different types of markets, fees, devices, etc.). The solution we propose is the use of newly developed reinforcement learning algorithms for jointly optimizing the environment and control policies. // Damien Ernst

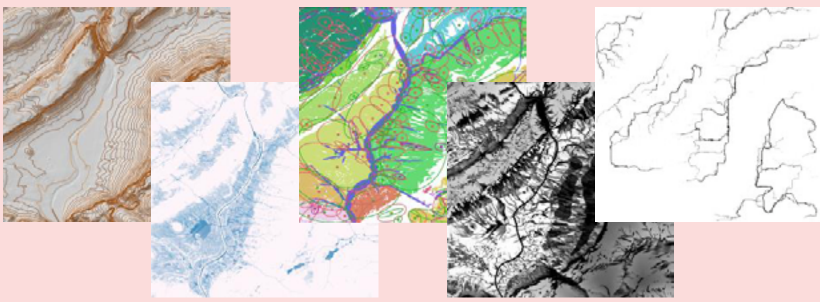
AI for good, in practice

Julien Grand-Clément is seeking to develop data-driven, interpretable policies for resource allocation in environments of scarcity (see next page). The themes of his research are aligned with the *United Nations Research Roadmap for COVID-19 Recovery*: How should health systems be designed so they are responsive, adaptable and accessible when needed? How can health systems eliminate discrimination in their service delivery and become drivers of equity in society? The guidelines he is elaborating are also contributing to the achievement of the *Good Health and Well-Being goal of UN Sustainable Development Goals*.

Visualizing climate change

Marie-Paule Cani's "smart 3D models provide quantitative validation" with existing ecosystems and a user study with expert paleontologist end-users, showing that our system enables them to author and compare different ecosystems illustrating climate changes over the same terrain while enabling relevant visual immersion in consistent landscapes.

▾ IMAGES EXTRACTED FROM
P. Ecormier-Nocca, G. Cordonnier, P. Carrez, A. Moigne, P. Memari, B. Benes, M.-P. Cani, *"Authoring Consistent Landscapes with Flora and Fauna,"* ACM Transactions on Graphics, Aug. 2021.



Data-driven, interpretable policies for resource allocation in environments of scarcity

The COVID-19 pandemic has exacerbated many existing inequalities in access to care. Allocation, and particularly rationing, of healthcare resources is a challenging and fraught decision that policymakers and providers may be forced to make during a pandemic, a natural disaster, or a mass casualty event. Well-defined guidelines to manage scarce life-saving resources must be designed to promote transparency, trust and consistency. To facilitate buy-in and use during high stress situations, these guidelines need to be interpretable and operational.

At **Hi! PARIS** we aim to develop novel data-driven **models** to guide such decisions. We use the allocation of critical care resources in the COVID-19 pandemic as a canonical setting to develop ethical, evidence-based, operational decision rules to guide resource allocation and mitigate imbalances in pandemic responses. Indeed, a major contributor to inequalities has been the geographical mismatch between demand for care and the availability of **healthcare** resources. In the U.S. and Canada, many smaller community hospitals close to where essential workers and minorities lived and worked were overwhelmed, whereas larger hospitals in high-income neighborhoods had the capacity to admit new patients.

Despite potential benefits, transferring patients among hospitals is a highly complex decision-making task. Patient transfers are costly and time-consuming for care providers with potential implications for patient safety and clinical outcomes in addition to an emotional toll for the patients families. Besides clinical considerations, the uncertainties in patients' length of stay, care requirements and the evolution of the pandemic add to the complexities of designing effective transfer policies that balance congestion among hospitals and reduce inequalities in access to care and its outcomes.

Once patients are admitted to a hospital, there are often questions on how to triage life-saving resources, such as ventilators, when supplies are limited. Triage guidelines are often implemented during high stress, complex situations. Therefore, triage algorithms need to be systematic, simple and intuitive in order to facilitate adoption and to ease the decision burden on the provider. On the one hand, government officials have issued pre-specified and transparent utilitarian triage guidelines for preventing loss of life, promoting fairness, and supporting front-line clinicians. Unfortunately, these guidelines are generally not constructed in a data-driven way, but rather via the expert opinions of clinicians, policymakers, and ethicists. Therefore, it is unknown how well they perform for the intended purpose of directing scarce resources to those most likely to benefit. In addition, it is not unethical to perform a prospective study to determine the efficacy (or performance) of such policies. Healthcare delivery operates in a resource limited environment where demand can sometimes exceed supply, resulting in situations where providers have to make difficult decisions about who and how to prioritize care. Having a framework to guide such decisions is critical, particularly with the growing threats of pandemics, natural disasters, and mass casualty events that make the healthcare system vulnerable to situations where demand vastly exceeds supply of critical healthcare resources.

This research involves a unique bi-national team who have extensive expertise in utilizing data-driven approaches to improve healthcare delivery as well as close clinical collaborators and clinical ethicists who will provide invaluable insights to ensure that the models are grounded in practice and that the recommended guidelines facilitate increased access to care while accounting for important ethical considerations. The interdisciplinary nature of this team, including members with backgrounds in operations, artificial intelligence and clinical matters, brings a novel perspective to pandemic management with a high potential for impact for policymakers, hospital administrators, and clinicians.

Julien Grand-Clément is Assistant Professor in the Information Systems and Operations Management Department. He conducts research on medical decisions automation using machine learning and optimization. He completed his PhD at the Industrial Engineering and Operations Research Department at Columbia University in 2021, and his MSc at École Polytechnique (Paris) in 2016.

His research has been published at international medical and artificial intelligence conferences, as well as in operations research journals. His latest collaborations include robust allocations of beds in intensive care units with hospitals in California and interpretable ventilator allocation guidelines for hospitals in New York City.

• **Hi! PARIS Chair holder**

Data-driven Models and Algorithms to Mitigate Imbalances via Resource Allocation for Pandemic Resilience

To develop data-driven, interpretable policies for resource allocation in environments of scarcity, we utilize a combination of econometric and statistical methodologies with stochastic modeling and algorithm design.

Stochastic models and algorithm design for resource allocation

A primary focus is to develop robust and interpretable policies in order to facilitate adoption in practice.

First, we incorporate more **flexibility** in decision criteria, by considering alternative review periods, for instance. This requires developing a novel methodology to not only learn what allocation decision to take but also *when* to take a decision (in contrast to taking a decision at each period of the stochastic process). Capacity-dependent triage protocols are also elaborated, the main challenge here being to overcome the curse of dimensionality with approximation methods.

Second, we develop robust triage policies thanks to a methodological approach incorporating parameter uncertainty and by designing triage algorithms that explicitly account for this uncertainty. We focus on the specific impact of unobserved confounders on the errors in parameter estimations, as they are likely to occur in a healthcare dataset. The rich econometric literature on unobserved confounders is of great help here in building confidence regions that specifically account for confounders in our sequential decision problem.

We then develop new stochastic network models that keep track of the health states of patients and account for the causal effect of patient transfers. We utilize the models to provide accurate predictions and inferences on future hospital occupancy levels and to develop "effective" transfer policies.

Data analytics and insights

We leverage unique data from different health systems during their COVID-19 response to develop data-driven insights.

Using clinical data, we will empirically measure the amount of parameter uncertainty in practice. The novelty of our approach is to take this uncertainty into account during the search for a better policy. This is in contrast with the classical approaches in the medical literature (e.g., sensitivity analysis), which first choose a policy and then study how the performances of the chosen policy change with a change in parameters.

The algorithmic approach developed previously helps us to develop triage algorithms and evaluate their potential performance. We expect these triage algorithms to improve upon current practices and official guidelines, which were designed by experts and policymakers but were not informed by the data. Our direct collaborations with hospital practitioners help us develop actionable and implementable guidelines.

Eventually, using tools from causal inference, we will estimate the impact of patient transfers on patient outcomes and other operational metrics, thus informing the models proposed previously.

Reference

E. Chuang, J. Grand-Clément, J.-T. Chen, C. W. Chan, V. Goyal, and M. N. Gong, *"Quantifying Utilitarian Outcomes to Inform Triage Ethics: Simulated Performance of a Ventilator Triage Protocol under Sars-CoV-2 Pandemic Surge Conditions.,"* AJOB Empir Bioeth, pp. 1–9, Apr. 2022.

AI & education

IP Paris and HEC Paris have created **Hi! PARIS** to develop education, research and innovation in the fields of artificial intelligence and data analytics. In this framework, considering the relationship between AI & education is a threefold task: (i) developing AI programs for students; (ii) examining how the training system is stressed by externalities, and what AI solutions are currently used and likely to shape its future; (iii) using AI tools to understand changes in learning practices, for instance in social learning. Here is how some of our Fellowship and Chair holders connect these questions to their research project.

Our goal is to push forward the frontiers of our understanding of the complex algorithms currently used in AI. This increased understanding will eventually have consequences on education, since an algorithm that is better understood is easier to teach. // Arnak Dalalyan

Creative AI allows students to manipulate certain matter, you can immerse them in our models describing geological phenomena, for example. As (Nobel laureate) Richard Feynman once wrote: 'What I cannot create, I do not understand.' // Marie-Paule Cani

Teaching AI

Several programs for students are organized by **Hi! PARIS**, responding to a call for more graduates with an expertise in data science and management, across all sectors of the economy. Here are some of the lectures and education courses in which the researchers presenting their work in these pages are involved.

Aymeric Dieuleveut gives lectures at most levels in connection to his Federated Learning research work: "*Collaborative and reliable learning*", "*Generalization Properties of learning algorithms*", "*Optimization and Deep learning*"

Klaus Miller teaches Artificial Intelligence in Management.

Anna Simoni teaches a class of "*Machine Learning and Big Data in Econometrics*" and is involved in a new course titled "*Machine learning and Macroeconometrics*".

Matthew Yeaton teaches "*Technology Strategy in the Age of AI.*"

Aluna Wang has designed an executive education course "*RegTech* (Regulatory technology)*: The Intelligent Risk Management*" partially based on her stream of research.

A training system under stress

The pandemic has stressed the need for an efficient training system to help workers adjust their skills in periods when there are strong asymmetric sectoral shocks on labor demand.

As vice-president of the scientific committee of the "*Plan d'Investissement dans les Compétences*", Roland Rathelot has followed how the 15-billion-euro program set up by the French government for 2018-2022 in order to upgrade the French training system targeting low-skill and unemployed workers has been spent.

Exploiting graph representation and the network structure of accounting data

The existing accounting literature focuses on investigating the determinants of accounting fraud and mainly deals with financial statement level data. Meanwhile, the data science literature contains two major strands on anomaly detection in financial transaction data. However, they use supervised or semi-supervised machine learning models. At **Hi! PARIS**, Aluna Wang, assistant professor at HEC Paris, uses the unsupervised machine learning model to address the spontaneity of anomalies and data imbalance and to build general, scalable, and explainable anomaly detection models. Meanwhile, she exploits graph representation and the network structure of accounting data to further improve the accuracy and robustness of anomaly detection.

This work also helps to illustrate how to use dynamic graph mining for specific risk management purposes such as AML, email communication monitoring, and fraud detection.

Unsupervised anomaly detection in financial transaction data

The Minimum Description Length (MDL) principle is a general theory of inductive inference—the basis of machine learning. It is particularly well-suited to tackling model selection, prediction, and estimation problems in situations where the models under consideration can be arbitrarily complex and where overfitting is a serious concern. We introduce pattern recognition and **anomaly detection** models for bookkeeping data based on the MDL principle. Unlike previous accounting studies, we use transaction-level data produced by the double-entry bookkeeping system. Furthermore, we employ **graph**-based techniques to capture the interdependencies of accounts and provide more adversarially robust fraud detection tools.

Using journal entry data from companies in different sectors, we show the effectiveness and robustness of our models in recognizing transaction patterns and identifying anomalies via convincing case studies and the successful recall of injected anomalies orchestrated by seasoned accounting practitioners. Our models outperform benchmark models by a significant margin. We demonstrate the value of combining domain knowledge and machine learning methods in graph construction and model building.

References

P. Liang, A. Wang, L. Akoglu, and C. Faloutsos, *"Pattern Recognition and Anomaly Detection in Bookkeeping Data."* (2021).

Dynamic graph mining for anti-money laundering, email communication monitoring, and anomaly detection

How can we spot money laundering in large-scale graph-like accounting datasets? How to identify the most suspicious period in a time-evolving accounting graph? What kinds of accounts and events should practitioners prioritize under time constraints? To tackle these crucial challenges in accounting and auditing tasks, we propose a flexible system called AutoAudit, which can be valuable to auditors and **risk** management professionals.

To sum up, there are four major advantages of the proposed system: (i) "Smurfing" Detection spots nearly 100% of injected money laundering transactions automatically in real-world datasets. (ii) Attention Routing attends to the most suspicious parts of time-evolving graphs and provides an intuitive interpretation. (iii) Insight Discovery identifies similar month-pair patterns proved by "success stories" and patterns following Power Laws in log-logistic scales. (iv) Scalability and Generality ensure AutoAudit scales linearly and can be easily extended to other real-world graph datasets. Experiments on various real-world datasets illustrate the effectiveness of our method.

M.-C. Lee, Y. Zhao, A. Wang, P. J. Liang, L. Akoglu, V. S. Tseng, et al., *"AutoAudit: Mining Accounting and Time-Evolving Graphs,"* in 2020 IEEE International Conference on Big Data (Big Data), 2020, pp. 950–956.

Knowledge spillovers in the open-source community: Evidence from GitHub networks

Knowledge spillovers are regarded as essential drivers for innovation. However, due to the invisibility of **knowledge transfer** and data limitation, few empirical investigations have been conducted to measure the impact of network formations on the effectiveness and efficiency of knowledge spillovers. This is a research topic Aluna Wang is working on at HEC Paris.

The difficulties are twofold: first, it is hard to quantify the knowledge transferred from one to the other; second, the network formation might also affect the efficacy of knowledge transfer. For instance, in the research co-authorship network, it is hard to measure to what extent knowledge about publishing papers in top journals could be transferred from established professors to their co-authors. A top publication as an outcome could be attributed either to the more established authors' reputational effect or to their substantial improvement of the paper's quality.

First, different behavioral patterns of developers have been identified and their contribution dynamics along with project life cycles have been studied. Second, after the appropriate data cleaning process, a graph database of two networks, i.e., the contributors' network and the project network, has been built. Hypotheses explaining contributors' motivation and contribution patterns based on theories in economics and information systems and channels through which network formations affect knowledge spillovers have been explored.

Finally, by applying advanced pattern matching algorithms, Aluna and her collaborators are now tracking the flows of source code pieces and building a general framework explaining how knowledge accumulates to successful projects and then spread to others, and how this process interacts with different network topologies.

Advantages of studying the open-source community

Community-based organizations are becoming increasingly important for product creation. It is claimed that knowledge spillovers enable other software developers and other projects to benefit from the innovation generated by a particular project in such scenarios. **Open-source** projects are semi-structured groups of talented programmers collaborating on interdependent tasks via informal, non-hierarchical, decentralized communication with the shared objective of developing a valuable product. The open-source innovation paradigm has advanced in recent years, as technology continues to make it simpler for developers to work remotely. Additionally, open-source development teams make the underlying project information available to the broader public, which eliminates many common intellectual property obstacles. One of the primary benefits of open-source development is the opportunity to exchange and absorb information developed outside of a single open-source project. Knowledge spillovers likely occur frequently in the open-source community because the source code is publicly available, and developers collaborate on various projects.

Application to other domains

The methodology used in this research can be applied to analyses of other industries in which knowledge spillovers play an important role, such as bio-pharmacy, academic co-authorship and social networking platforms. For instance, a collaborative relationship among firms can be regarded as a knowledge transmission network. A tentative mergers and acquisitions deal could change the formation of the network and thus have far-reaching implications for the underlying knowledge flows and industrial productivity.

• **Hi! PARIS Chair holder**

Accounting, AI and Cybersecurity

Aluna Wang received her PhD and MSc in Industrial Administration from Carnegie Mellon University, where she used to be PwC Presidential Fellow at the Digital Transformation and Innovation Center. Her research endeavors feature two major themes. One theme is to examine how information transmission mechanisms in the financial market, such as public disclosures and lending relationships, affect real economic outcomes. Aluna strives to provide insights into how regulating information in the financial markets shapes the real economy. The other is to develop and deploy machine learning-based tools to improve our understanding of accounting data and provide intelligent solutions to the real-world challenges faced by financial service professionals in our rapidly changing digital landscape.

Culture, language, and networks
in social communities

Exploring how digitization informs and develops our understanding of organizational culture, knowledge transfer, and the labor market provides some welcome insights. Indeed, digitization has opened a window into network structure and language, providing a lasting record of these changes over time. Using these digital records to observe the structure of social relations and the language used to communicate can help deepen our theory of knowledge transfer for a wide range of organizations, not just those that operate in the digital sphere. This means that these studies also have implications for our understanding of organizations in non-digital settings.

Effect of communication networks on cultural change in organizations

Social networks and organizational culture are deeply connected. However, the causal direction between the two is a thorny problem. To what extent and through what mechanisms can social network structure drive cultural and linguistic change?

At **Hi! PARIS**, Matthew Yeaton (see page 16) adopts a two-pronged approach to this question: a formal model to help unpack the mechanisms, and an empirical test using a natural experiment with an Alt-Right online community on the Reddit platform. The model first develops how the culture of an organization, including an online community, is reflected in the language of those who belong to that organization. After incorporating the dynamic interplay between social network structure and language in the presence of **social learning**, it is suggested that language is partially endogenous to network structure. That is, an organization can shock or "rewire" the network to change users' language. Moreover, any such shock must also increase **conductance**, a measure of the relative insularity between individuals using idiosyncratic language and others. In the context of a shock to the social network, changes in conductance reflect changes in the structural cohesion of interconnected groups. In other words, increasing the conductance between two groups means we are making the groups more structurally integrated. This structural integration, captured by increasing conductance, illustrates that a key mechanism in changing language is the increase in social learning across these groups.

All the above has been tested using a natural experiment with the aforementioned Alt-Right online community. The natural experiment shocks the network, and linguistic change following this shock is evaluated thereafter. This experiment shows causal evidence of the endogeneity of language to social network structure. Across several ways of measuring language use, including measures derived from contemporary advances in natural language processing techniques, we find that this shock causes a shift in language use away from the jargon of the Alt-Right community (including a decrease in hate speech).

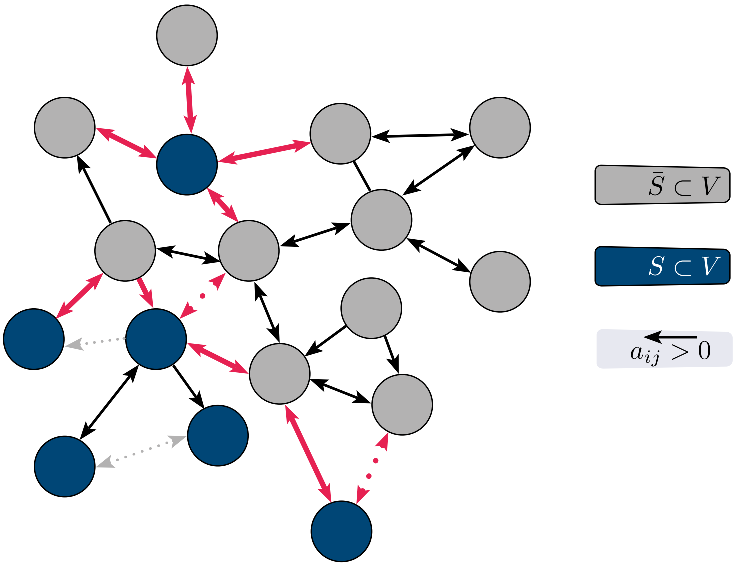
The Organizational Ship of Theseus

In a collaborative work with an organizational theorist*, we investigate the ways that culture, and especially language, can encode organizational priorities and enable organizational memory. Culture's ability to encode priorities creates differences between organizations in which types of knowledge can be efficiently communicated. Language in this formulation has a connection to routines: they are held at the organizational level, and can be a means to transfer knowledge efficiently within an organization, consistent with the knowledge-based view of the firm. Language also takes on an emergent property, that of storage. This is because we cannot use language to communicate without also communicating its priorities via its structure.

Like an organizational Ship of Theseus, we ask how much of the language of the original members remains when they have all been replaced. We find that by iteratively replacing every member of the or-

▲ *The Ship of Theseus (or Theseus' paradox) is a thought experiment related by Plutarch in "Life of Theseus". It raises the philosophical question: does an object that has had all of its components replaced remain fundamentally the same object?*

* partnerships are listed at the end of the document



◂ *Directed network emphasizing the conductance-relevant edges*

For the conductance between two disjoint groups S (blue) and $\bar{S}$ (grey) to increase, we could increase the amount of influence from existing edges (making the solid red lines larger), we could create new "across the aisle" edges (the dotted red lines), or, as is most relevant to our empirical context, we could decrease the influence of existing edges within a group (making the dotted light grey lines smaller).

ganization, a substantial percentage of the updated language reflects the founding members' priorities. Moreover, we find that the problem is exacerbated for larger organizations. Even after two full replacements of the organization—everyone has been replaced, and their replacements have been replaced—about 15% of the updated language reflects the original members' priorities.

Common language is a type of limited codifiability that makes certain (possibly tacit) routines and situations more efficient to describe within an organization. Thus, the upshot of common language is that it provides a pathway for organizational memory. On the other hand, this also suggests that it is not enough to change the language and culture of an organization simply by changing its members, and thus that this may be one part of the puzzle of organizational rigidity.

We also test the model using data from a popular online strategic communication cooperation game. By exploiting a panel of partially overlapping teams, we find that players bring communicative routines with them when they join a new team, and are able to transfer this knowledge to the new team.

Leveraging the setting of digitization as a window into the culture and networks of social communities

Identifying exiting and voicing behaviors

In another collaborative work, we revisit Hirschman's (1970) "Exit, Voice, and Loyalty" theory in the context of the gig economy, leveraging **text analysis** to analyze data from the largest online forum for Lyft and Uber drivers. Workers who are dissatisfied with an organization may express their discontent by exiting or by employing their voice, options which are traditionally treated as substitutes. As outside options become more attractive, exit rises and voice falls. In updating the framework, we introduce the idea of partial exit—gig workers need not exit a platform entirely when dissatisfied, but rather can adjust their allocation of labor between platforms—and discuss the imbalanced power relationship between gig workers and employment platforms.

We argue that under the conditions of platform gig employment, both exit and voice are likely to rise as alternative options improve. We empirically test our predictions in the ride-sharing market: exploiting time variation in Lyft's market share gains on Uber in 59 U.S. cities between 2014-2018, we analyzed posts from the largest online forum for ride-sharing drivers. Analyzing conversations at the "digital water cooler" allows us to quantify how drivers' discussions of the exercise of both exit and voice shift in response to market conditions. We show that as Lyft gained market share, drivers in those cities both increased their discussion of signing up and working for Lyft (partial exit), and increased their discussion of labor organizing (voice). Using topic modeling, we provide detailed evidence of the specific conversational topics that shifted in response to Lyft's market share gains.

Federated Learning
a multi-faceted key challenge

Over the last five years, a new domain in machine learning has emerged and gained significantly in importance, as a fundamental setting to tackle new societal and industrial challenges: **Federated Learning**.

In Federated Learning, several organizations or devices seek to collaboratively train a machine learning model under the orchestration of a central server, while keeping individual datasets on their respective local storage, in other words, *without* centralizing data.

Due to this decentralization principle, Federated Learning is often associated with the quest of data protection. Indeed, privacy has become a major concern for both society (individuals participating in training want to protect their privacy) and industries (which face legal constraints preventing them from exploiting valuable data). However, other needs and constraints have found a path toward a solution in such learning schemes. Overall, it has become necessary to train the models without centralizing the data, either because of those privacy constraints, or because extremely large datasets have to be distributed over networks of storing devices.

Four challenges ahead

With this new form of distributed learning, new optimization challenges are arising, taking into account communication constraints, while new opportunities to improve the models to adapt to users are being explored. Developing new algorithms for Federated Learning is thus a key challenge. This will simultaneously positively impact society, by protecting individual data and restoring public trust and confidence in machine learning technologies, and unlock countless novel opportunities for partnerships between entities willing to collaborate without centralizing their datasets, a crucial factor in medical applications, fraud detection, Internet of things, and many other domains.

At **Hi! PARIS** we identify and consider four pivotal challenges of Federated Learning: **communication constraints**, **statistical heterogeneity**, **missing data**, and **privacy**.

In our research directions we can tackle *a combination* of several of those challenges, as shown overleaf.

Our goal is to build the next generation of algorithms for large-scale Federated Learning, supported by strong theoretical guarantees, and practical implementations together with open-source code. These new algorithms, methods and guarantees constitute a key step in building reliable Federated Learning architectures, and are expected to bring major advances to the field by improving the scalability and efficiency of methods, and enabling fruitful collaborations that would not have occurred otherwise.

To do so, we rely on a precise mathematical understanding of optimization algorithms and of the statistical trade-offs of large-scale learning.

Aymeric Dieuleveut is an Assistant Professor in Statistics in the applied mathematics department at École Polytechnique. His main research interests are statistics, optimization, stochastic approximation, Federated Learning, high-dimensional learning, non-parametric statistics, and scalable kernel methods.

He was a postdoctoral researcher at EPFL, Lausanne, in the MLO team, directed by Martin Jaggi. He was also a PhD student in the Sierra Team, which is part of the DI/ENS (Computer Science Department of École Normale Supérieure), under the supervision of Francis Bach.

http://www.cmap.polytechnique.fr/~aymeric.dieuleveut/

• **Hi! PARIS Fellowship Starting**

FLAG- Federated Learning: new Algorithms with Guarantees

Federated Learning is not solely a matter of privacy

Studying missing data in linear models

Missing values arise in most real-world data sets due to the aggregation of multiple sources and intrinsically missing information (sensor failure, unanswered questions in surveys, etc.). In linear models, in the presence of missing values, the Bayes rule can be decomposed as a sum of predictors corresponding to each missing pattern. This eventually requires solving a number of learning tasks that are exponential in the number of input features, which makes predictions impossible for current real-world datasets.

We propose a new thresholded predictor with strong theoretical guarantees: the upper bound on its excess risk holds under very mild assumptions on the data, while integrating a quantity that describes the influence of the missing data distribution on predictive performances.

REFERENCE

A. Ayme, C. Boyer, A. Dieuleveut, and E. Scornet, *"Minimax rate of consistency for linear models with missing values,"* accepted at International Conference on Machine Learning (ICML 2022).

Tackling communication constraints

We propose MCM, a new algorithm for bi-directional compression in Federated Learning with **heterogeneity** constraints and fast convergence. The downlink compression only impacts local models, while the global model is preserved. As a result, the gradients on local servers are computed on perturbed models. The precise control of this perturbation is ensured in MCM by the combination of model compression and a memory mechanism.

We demonstrate that the server can drastically reduce its communication burden without impacting convergence of the algorithm. This is promising, for instance to reduce the energy impact of algorithm training. This work also opens new doors, e.g. incorporating worker-dependent randomized-models and partial participation.

REFERENCE

C. Philippenko and A. Dieuleveut, *"Preserved central model for faster bidirectional compression in distributed settings,"* in Advances in Neural Information Processing Systems, 2021, vol. 34, pp. 2387–2399.

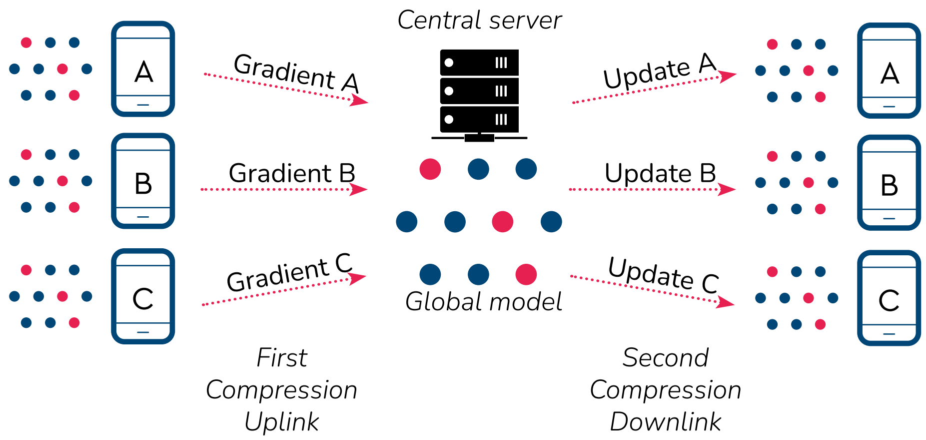
Differential privacy with data heterogeneity and communication constraints under strong theoretical guarantees

We propose a new algorithm, DP-SCAFFOLD, that ensures the confidentiality of the agent's data while being robust to the statistical heterogeneity of workers. We focus on a challenging setting where users communicate with a "honest-but-curious" server without any trusted intermediary, which requires ensuring **privacy** not only toward a third party observing the final model but also toward the server itself.

REFERENCE

M. Noble, A. Bellet, and A. Dieuleveut, *"Differentially Private Federated Learning on Heterogeneous Data,"* International Conference on Artificial Intelligence and Statistics, (AISTATS, 2022).

Illustration of the bidirectional compression framework with "model compression". At step 1, we send a gradient difference, at step 2 we communicate a model difference.



Designing models
new and effective ways

Model complexity constraining

The broad success of DNNs stems from their ability to automatically learn from examples, in general without requiring any specific expertise, and using very general learning strategies. The use of GPUs for training DNNs boosted their large-scale deployability; however, there are some obstacles toward the universal embeddability of DNNs in many IoT devices.

Indeed, typical AI architectures require a lot of memory and a lot of computation. Focusing on the latter aspect, one common way to estimate the complexity of these operations is to measure the required floating-point operations (FLOPS). As a practical example, in order to process a 224x224 pixels image, the ResNet-50 architecture requires 7.7 GFLOPs, while DeepLabv3—for image segmentation—requires 51 GFLOPS.

As the research around AI has progressed in the last decade, Deep Neural Networks (DNN) have been incrementally modified in order to achieve high performances for the specific task they have been trained to solve (in this case, high performance means high accuracy in their prediction, or in the output). However, the AI efficiency aspects (namely, required memory, number of operations, power consumption) have in general been marginal. Even in recent years, the growing complexity of models like DenseNets and Transformers has challenged the computational capability of embedded devices, which are required to be more and more powerful to be able to execute modern architectures.

The complexity of DNNs can be reduced by enforcing a **sparse network topology**, that is, some connections between neurons can be pruned by wiring the corresponding parameters to zero. Besides the reduction in parameters, some works also suggest other benefits derived from pruning DNNs, like improved performance in transfer learning scenarios.

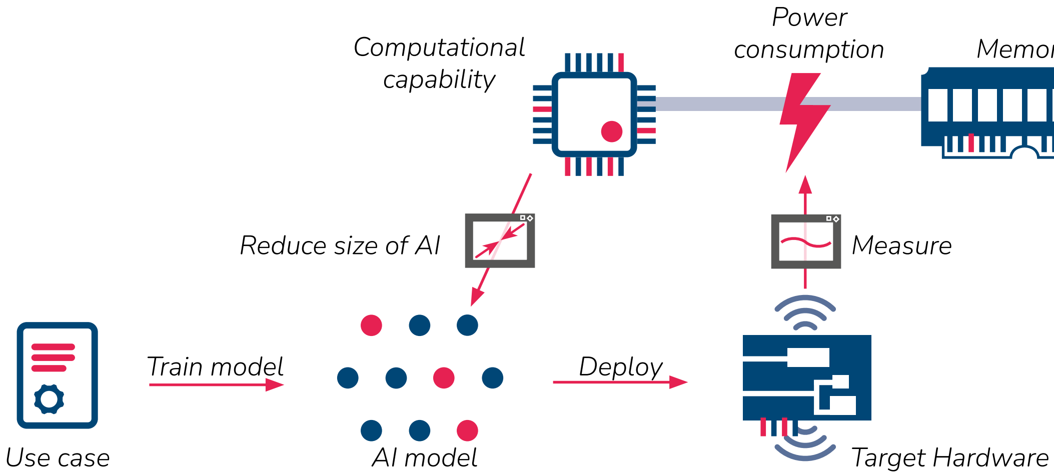
Popular methods introduce *a regularization term* in the cost function with the goal of shrinking some parameters to zero. Next, a threshold operator pinpoints the shrunk parameters, eventually enforcing the sought sparse topology. However, such a method requires that the topology to be pruned has been preliminarily pruned via standard gradient descent, which sums up to the total learning time.

Refining regularization strategies

A few years ago, a regularization strategy which penalizes all the parameters having little or no impact on the generation of the output was proposed.

This measure is defined as sensitivity S of the output of the DNN model with respect to a parameter of the model. It quantifies how the output of a trained model changes for small variations of the parameter. Let us say that the model has been already trained: if the output of the model is insensitive to these small variations, then the parameter is not important to the generation of the output, and can be removed from the model; otherwise, the output will be sensitive to small variations of the parameter for the given task, and the parameter cannot be modified, hence, removed.

As a follow-up improvement, a sensitivity-based technique referring to the loss function (namely LOBSTER) has been also developed. As it refers to the loss value and no longer to the raw output of the model, it can be successfully deployed directly at training time on a non pre-trained model. Additionally, we have developed a technique which removes entire neurons instead of single parameters (SeReNe). Working at the entire neuron level rather than at that of the single parameter is more challenging and concretely enables a higher memory footprint reduction as it leads to a layer's rank reduction. We have also shown within the *MPEG 7 part 17 pipeline* that, despite removing fewer parameters, structured sparsity enables higher model **compressibility**.



Deep Neural Networks efficiency scheme

Designing efficient and energy-friendly models

The design of an effective optimization strategy for the AI model's energy and memory efficiency will be one core challenge for the near future. Unlike with state-of-the-art approaches, it will be key to address the problem of designing a good estimator of the hardware's real memory and **energy consumption** rather than blindly minimizing the model's parameters and/or neurons. Raw evaluation and estimates of the model's complexity using a crude proxy like FLOPS evaluation are considered as suboptimal after evaluations in the field. Indeed, it is intuitive that factors like data transfer costs—hence, caching, RAM availability, cores—impact power consumptions in the hardware. Indeed, state-of-the-art power estimation approaches have been shown to be very limited in predictions.

Our objective is to achieve efficient DNN models, while still satisfying the target performance expected on the specific task. This efficiency can be expressed as a function of a number of parameters (synaptic connections) of the model, memory occupation, and FLOPS, aiming at estimating the final power consumption of the deployed model. Such a constraint should adapt to the different hardware resources the ANN is deployed to, depending on the specific hardware and on specific applications, with a scheme like the one proposed in the figure above.

Focusing on another aspect of **frugal AI**, namely training efficiency, recent advances in deep learning optimization show that just a subset of parameters are necessary to successfully train a model. Potentially, this discovery has a broad impact, from theory to application; however, it is known that finding these trainable sub-networks is typically a costly process. We explore this possibility in order to understand why common approaches typically fail in the extreme scenarios of interest, and prospectively propose an approach which potentially enables training with a higher energy efficiency.

References

E. Tartaglione, S. Lepsøy, A. Fiandrotti, and G. Francini, *"Learning Sparse Neural Networks via Sensitivity-Driven Regularization,"* in Proceedings of the 32nd International Conference on Neural Information Processing Systems, Red Hook, NY, USA, 2018, pp. 3882–3892.

E. Tartaglione, A. Bragagnolo, A. Fiandrotti, and M. Grangetto, *"LOss-Based SensiTivity REgulaRization: Towards Deep Sparse Neural Networks,"* Neural Netw., vol. 146, no. C, pp. 230–237, Feb. 2022. See also: https://github.com/EIDOSlab/LOBSTER

A. Bragagnolo, E. Tartaglione, A. Fiandrotti, and M. Grangetto, *"On the Role of Structured Pruning for Neural Network Compression,"* in 2021 IEEE International Conference on Image Processing (ICIP), 2021, pp. 3527–3531.

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

E. Tartaglione, S. Lathuilière, A. Fiandrotti, M. Cagnazzo, and M. Grangetto, *"HEMP: High-order entropy minimization for neural network compression,"* Neurocomputing, vol. 461, pp. 244–253, Oct. 2021.

Using and producing data with care

Research conducted at **Hi! PARIS** is both theoretical and applied, with each side nourishing the other. Datasets and the algorithms that manipulate data are not just mathematical objects to study. They often concern the daily life of people, whose fundamental rights must be preserved. The robustness and reliability of AI solutions have to be guaranteed by theoretical works, confronted with the ground truth on a regular basis.

Some datasets used

Roland Rathelot and his collaborators were the first team to build a dataset to follow the full job search activity of unemployed job seekers, finding out their demographic characteristics, as well as the firm in which they find a job and their reemployment earnings (see page 15). They use this data to study gender gaps in job searches. In particular, they ask whether women and men click on and apply for jobs with the same intensity, what are the characteristics of the jobs they apply for, and how this relates to the jobs they end up with. In a previous work using French data, they had shown that gender gaps in commuting preferences were a modest yet significant factor of gender wage gaps.

To study the dynamic evolvement of developers' contributions and the underlying knowledge flows, Aluna Wang exploits data from GitHub, the world's largest open-source development network (see page 33).

The available datasets at the beginning of the pandemic were heavily biased

Where there are high stakes, there is a need for guaranties

The adoption of machine learning in high-stakes industrial application areas calls for strong guarantees: is the learned model accurate enough? How is fairness taken into account, during training time and during inference? Are the models or algorithms vulnerable to adversarial attacks? How are the resultant data protected against leaking and divulgation, from training time to inference?

Answering these questions requires dealing with hard-to-control and even unknown quantities—such as risk and regret—and various and not always technical constraints—such as privacy, fairness, robustness, reliability, and transparency. These quantities and complex constraints are manipulated in the learning algorithms as a search for optimized bounds. Major fundamental mathematical works must be carried out, with the twofold benefit of providing guarantees and helping to open the ML black-box.

Reliability and robustness

As it is increasingly common to generate synthetic data to feed AI models, the question is raised whether these datasets are still guaranteed to be reliable.

Arnak Dalalyan works on generative models (see next page) which will help develop mathematical tools that are tailored to **highly parameterized and high-dimensional models**.

More precisely, he focuses on getting finite-sample statistical guarantees which will clearly highlight the impact of a suitably defined notion of intrinsic dimension (as opposed to ambient dimension) on the risk.

These kinds of guarantees are paramount in machine learning to ensure the reliability of the algorithms obtained. Furthermore, special attention is paid to stability and robustness properties: robustness to the model mis-specification and to the presence of **outliers in data**.

Statistical methods are omnipresent in the most powerful artificial intelligence technologies. The algorithms that are currently used rely on high-dimensional and highly overparameterized models (such as deep neural networks) for which classical theoretical results do not lead to a good understanding of the empirically observed phenomena.

Analysis of generative models

At **Hi! PARIS** we carry out fundamental research highlighting the properties and the limitations of different algorithms currently used in many applications.

The research work conducted by Arnak Dalalyan, professor at ENSAE Paris, focuses on **federated learning** and **optimization**. Indeed, the task of **generative modeling** and **sampling from complex distributions** is intimately linked to the problem of optimization and is often carried out for large scale data. Taking advantage of distributed algorithms might result in considerable savings in computation, time and energy.

The broad goal of this research is to contribute to a better understanding of the mathematics underlying the spectacular results obtained by deep learning methods. To this end, we will develop mathematical tools that are tailored to high-dimensional and highly parameterized models by focusing on model averaging, approximate sampling, generative modeling and robustness.

While most recent approaches that justify the success of deep learning in theoretically terms seek to quantify the performance of the trained model as a solution to an optimization problem, we intend to study this question through the lens of model averaging and sampling from a given distribution. We also intend to investigate optimality properties in generative models such as conditional generative models and cycle-generative models.

We will first find guarantees for predictors constructed as averaged neural networks, before investigating the finite sample properties of generative models inspired by conditional-GANs and cycle-GANs. This will lead to an assessment of **robustness** in generative modeling, and will eventually help to design and analyze accelerated methods for sampling in highly parametrized models.

This research work is mainly mathematical, but it also has an important computer science component. Indeed, most if not all mathematical findings need to be checked and illustrated by numerical experiments. Furthermore, the computational feasibility of proposed new algorithms needs to be demonstrated.

Arnak Dalalyan received his PhD from Le Mans University in 2001 and his *Habilitation à Diriger des Recherches* from Paris 6 University in 2007. He is currently professor of Statistics and Machine Learning at ENSAE Paris, Institut Polytechnique de Paris. Since 2020, Arnak is the director of the CREST (Center of Research in Economics and Statistics). His research focuses on the mathematical foundations of statistics and learning: nonparametric estimation, dimension reduction, robustness and sampling. Arnak serves as Associate Editor of the journals Annals of Statistics and Bernoulli. He also regularly serves as area chair in machine learning conferences NeurIPS, ICML, COLT, ALT.

Director of CREST (2020-), Associate Editor of Annals of Statistics (2019-), Journal of Machine Learning Research (2019-), Bernoulli (2019-), Electronic journal of Statistics (2016-), Head of the Statistical Learning program at ENSAE Paris (2014-).

- **Hi! PARIS Fellowship Advanced**

SAGMOS : Statistical Analysis of Generative Models: Sampling Guarantees and Robustness

Using generative modeling to study high-dimensional probabilistic models

Studying **complex systems** made up of many interacting components is a major challenge arising across fields. A biologist and a chemist may be interested in a protein configuration made of hundreds of atoms surrounded by many more solvent molecules. An economist may want to predict the effect of a certain policy on a population of firms and individuals.

Studying the fine details of the behavior of such systems is typically not necessary; understanding their typical behavior via probabilistic modeling usually reveals the most important information. For instance, the average time fraction a protein spends in a given configuration or the expected unemployment rate following the policy. However, these high-dimensional probabilistic models remain notoriously hard to probe.

Generative modeling aims at the unsupervised learning of a probabilistic model able to generate data matching typical realizations provided as training data. A typical example would be the generation of an image with a particular content, say a face, starting from a collection of such images. A variety of approaches has now proven that deep learning-based generative models can generate impressively faithful images, audio or texts. The question is now whether deep generative models can be leveraged to study high-dimensional probabilistic models.

> *Deep insights from complex systems can be drawn from a probabilistic point of view combined with statistical inference*

Powerful and close to universal methods to study complex probabilistic models are **Monte Carlo methods**, which simulate typical realizations of the model—or sample from it—to infer its properties. Despite the wide applicability of Monte Carlo methods, they can become arbitrarily costly in high dimension for distributions with complex geometries or metastabilites. In this context, deep generative modeling is a promising tool to assist **sampling**.

However, there are major differences between generative modeling and sampling from a given distribution associated with a probabilistic model. The first uses a large amount of data to create a model able to generate easily matching data, while the second seeks to obtain data realizations from the mere knowledge of a model. The measure of success also differs between the two programs. For sampling, the objective is well defined and performances can be assessed objectively on chosen examples. On the other hand, generative modeling is an ill-posed problem and subjective cues are sometimes the best way to evaluate the success of a model, such as the quality of an image. As a result, a major question is whether generative modeling can be precise enough to meet the requirements of successful sampling.

Marylou Gabrié obtained her PhD in Physics from École Normale Supérieure in 2019, after which she moved to New York for a shared young faculty position between the Center for Data Science of New York University and the Center for Computational Mathematics of the Simons Institute.

She received the L'Oréal Fellowship for Women in Science in 2018.

• **Hi! PARIS Chair holder**

High-dimensional probabilistic modeling and inference at the crossroads of machine learning and statistical mechanics

The intersection between generative modeling and sampling has been an active research area for a few years: a first picture of points of difficulties has emerged. A series of these works inspired by Variational Inference has shown the potential of the strategy, yet only in relatively simple models. Other works have identified the need for samples to be available a priori for training to tackle more sophisticated systems. A strategy currently pushed at **Hi! PARIS** by Marylou Gabrié resolves this tension by employing an adaptive Monte Carlo strategy.

Adaptive Markov chain Monte Carlo

In a joint research work, we revisited the idea of **adaptivity**, which consists of fine-tuning Markov chain Monte Carlo (MCMC) kernels based on previous states visited by the chains. In our work, we leverage the expressiveness of **normalizing flows**, a type of generative model that allows tractable marginal likelihood estimation. Our work demonstrated that no representative sample of data was needed a priori to successfully speed up the convergence of the MCMC thanks to the normalizing flows, while we clarified the requirement to have prior knowledge of the rough location of modes of interest.

In a subsequent collaboration with Eric Moulines, we are pursuing this renewed investigation of adaptive MCMCs, deepening our understanding of their limitations and opportunities in overcoming them.

Reference

M. Gabrié, G. M. Rotskoff, and E. Vanden-Eijnden, *"Adaptive Monte Carlo augmented with normalizing flows,"* Proceedings of the National Academy of Sciences, vol. 119, no. 10, p. e2109420119, 2022.

Structure-aware generative models

In previous works along with collaborators, we identified the need for well-chosen parametrizations of normalizing flows to beat the curse of dimensionality. As convolutional neural networks have been the corner stone of automatic image processing, tailoring models to different data types is a pre-requisite for designing algorithms that scale well. This line of work was popularized under the phrase "geometric deep learning".

In a joint research work, we are currently using Score-Based Generative Models adapted to physical systems including elementary particle clusters. Other collaborations help us to develop fundamental methods to model physical systems.

*Most of these works rely on a specific type of generative model called **normalizing flows***

Numerical experiments show how scaling to high dimensions is also facilitated by parametrizations of normalizing flows that incorporate known structures of the target distributions, such as short-scale correlations. Figure "Detecting Phase Transitions in Interacting Particle Systems", extracted from "Adaptive Monte Carlo augmented with normalizing flows".

Applications across scientific fields

Combining outputs of Markov chains Monte Carlo (MCMC) and generative models, we are able to design algorithms matching the needs of a variety of scientific communities exploiting **high-dimensional** probabilistic models. Here are some examples:

- *Bayesian inference—in particular in the astrophysical context.* Bayesian inference is a popular choice when little data is available to fit a model to observations, as it allows one to incorporate priors and estimate uncertainties. The needs of the astrophysics community have been identified and adapted solutions proposed. A user-ready Python package that will make deep learning enhanced sampling widely accessible for this community is planned to be released.
- *Molecular systems—another field with intensive needs of efficient simulation algorithms is molecular biology.* Despite our understanding of fundamental interactions, the computational cost of *ab initio* simulations of fully resolved trajectories is prohibitive for anything else than very elementary molecules.
- *Data-driven probabilistic modeling—in particular in the shape of Energy Based Models (EBM).* Performing density estimation from data is a very generic task for which a variety of methods are available. Among them, fitting an EBM is a particularly versatile approach served by the universal approximation power of deep neural networks. Yet the main difficulty in training and exploiting EBMs lies in sampling from the learned model. A new strategy for EBM learning combining an adaptive MCMC assisted by a normalizing flow is currently under development. By solving the sampling problem of EBMs, this line of work will open new possibilities in data-driven probabilistic modeling.

AI for Sound

a hybrid approach with dedicated generative networks

AI for Sound is defined as the general field of Artificial Intelligence applied to **audio analysis, understanding and synthesis** by a machine. As in other domains of AI, the field has rapidly moved toward so-called end-to-end neural approaches, which aim to directly solve the machine learning problem for raw acoustic signals but often only loosely taking into account the nature and structure of the processed data.

At **Hi! PARIS**, Gaël Richard, professor at Télécom Paris, believes that our prior knowledge about the nature of the processed data, their generation process and their perception by humans should be explicitly exploited in neural-based machine learning frameworks to be able to build more **frugal** and **interpretable** AI systems. His main interest is therefore to build **hybrid** deep approaches combining parameter-efficient and interpretable signal models, musicological and physics-based models, with dedicated neural generative networks.

This research is at the heart of his ERC advanced project "HI-Audio". The main targeted applications include speech and audio scene analysis, music information retrieval, and sound transformation and synthesis.

REFERENCE

ERC Advanced project no101052978 ERC-2021-ADG (2022) *"HI-Audio (Hybrid and Interpretable Deep neural audio machines)"*, funded by the European Union.

Researchers at **Hi! PARIS** engage in numerous international academic collaborations across several disciplines. For the research works presented in this book, the Fellowships and Chair holders want to thank their colleagues from:

Cape Town University (South Africa), Carnegie Mellon University (USA), Chinese Academy of Sciences (China), CNRS (France), Columbia Business School (USA), ENS Paris (France), Flatiron Institute (USA), Goethe University Frankfurt (Germany), Institut Louis Bachelier (France), Institute for International Economic Studies (Sweden), KAUST (Saudi Arabia), London School of Economics and Political Science (UK), MIT (USA), National Yang Ming Chiao Tung University (China), New York University (USA), Paris School of Economics (France), Sciences Po Paris (France), Skema Business School (France), Stanford University (USA), Technische Universität München (Germany), Toulouse School of Economics (France), Twente University (Netherlands), Universidad de Antioquia (Colombia), Università Bocconi Milano (Italy), Università di Genova (Italy), University of Edinburgh (UK), University of Florida (USA), University of Liège (Belgium), University of Manchester (UK), University of Toronto (Canada), and Uppsala University (Sweden).

Hi! PARIS, an interdisciplinary center for research and education devoted to AI and Data Science, designed to better serve the interests of Science, Economy, and Society

The founders

The **Institut Polytechnique de Paris (IP Paris)** is a public higher education and research institution that brings together five prestigious French engineering schools: École Polytechnique, ENSTA Paris, ENSAE Paris, Télécom Paris and Télécom SudParis. Under the umbrella of the Institute, these schools combine two centuries of expertise in the pursuit of three major goals: excellence in education, cutting-edge research, and promotion of innovation. Thanks to the academic foundations of its five founding schools and its alliance with HEC Paris, IP Paris is positioned as a leading academic and research institution, both in France and internationally.

HEC Paris is specialized in education and research in management sciences. HEC Paris offers a complete and unique range of academic programs for the leaders of tomorrow. Founded in 1881 by the Paris Chamber of Commerce and Industry, HEC Paris has a full-time faculty of 140 professors, 4,500 students and 8,000 managers in executive education programs every year. Ranked among the best business schools in the world, with a student population from over 100 countries (constituting 40% of the total student population), HEC Paris aims to create a new model of business school for the 21st century.

In order to develop ambitious and long-term research projects, it is necessary to design a model of citizen patronage favoring the general interest of all, on the Anglo-Saxon model. In the framework of **Hi! PARIS**, the initiative has been taken of developing a new concept of patronage: the success of this project requires a break from the existing model. Six corporate donors: **L'Oréal**, **Capgemini**, **TotalEnergies**, **Kering**, **Rexel** and **Vinci** contribute to the evolution, alongside the Center, of today's French patronage model. These French flagships with worldwide influence, which have long supported research and development, are committed to helping France up its scale. Without their support and funding, this new Center could not have been established. It is thanks to them, and to the other French and European corporate donors who will join them, that research and teaching activities will be strengthened in order to increase France's level of competitiveness on this fundamental and priority theme.

New member

In addition to IP Paris and HEC Paris, other institutions are keen to contribute to the **Hi! PARIS** ambition. In July 2021, Inria joined forces with **Hi! PARIS**.

Inria is the French national research institute for digital science and technology. World-class research, technological innovation and entrepreneurial risk are its DNA. In its 200 project teams, most of which are shared with major research universities, more than 3,500 researchers and engineers explore new paths, often in an interdisciplinary manner and in collaboration with industrial partners to meet ambitious challenges. As a technological institute, Inria supports the diversity of innovation pathways: from open-source software publishing to the creation of technological startups (Deeptech).

Central to the **Hi! PARIS** governance, the International Scientific Advisory Board gathers 10 top scientists with recognized expertise in the research fields covered by the center.

Francis Bach, Inria – Kathleen Carley, Carnegie Mellon University – Anindya Ghose, New York University – Avi Goldfarb, University of Toronto – Michael Jordan, University of California, Berkeley – Roni Michaely, University of Geneva – Masashi Sugiyama, University of Tokyo – Mariarosaria Taddeo, University of Oxford – Mihaela van der Schaar, Cambridge University – Lenka Zdeborova, École Polytechnique Fédérale de Lausanne

Hi! PARIS thanks its six corporate donors:
L'Oréal, Capgemini, TotalEnergies, Kering, Rexel, and Vinci.
.

Contact

Executive Director: Pr. Gaël RICHARD

Email: contact@hi-paris.fr

Phone: +33 (0)1 75 31 96 60

https://www.hi-paris.fr/