

### Boosting digital transformation

# On some research and innovation challenges in Al

FRANCE

### Faïcel Chamroukhi, October 25, 2024.



**Trinity College Dublin** Coláiste na Tríonóide, Baile Átha Cliath The University of Dublin



www.irt-systemx.fr/en

Some content in this PDF may not have been accurately converted from the original PPTX presentation



System×

Challenges and industrial applications in

- Hybrid AI: How to exploit industry knowledge of physical (scientific) and symbolic (semantic) nature in data-driven learning models.
- Trustworthy AI: Towards integrating AI into critical system engineering
- Generative AI for industry: How to evaluate and advance specialized generative AI models for industrial applications



System×

Challenges and industrial applications in

- Hybrid AI: How to exploit industry knowledge of physical (scientific) and symbolic (semantic) nature in data-driven learning models.
- Trustworthy AI: Towards integrating AI into critical system engineering
- Generative AI for industry: How to evaluate and advance specialized generative AI models for industrial applications



Related to the design and supervision of complex (physical) systems

- Covering various fields in physics (mechanics, fluid dynamics, aerodynamics, electromagnetism ...)
- In a wide variety of Applications in industry, in particular in numerical simulation



Picture from Marot, A., et al. (2018).

Picture form Merino-Martínez et al. CEAS Aeronautical Journal (2019).

From HSA – SystemX

from Emmanuel Menier (PhD, LSIN/SystemX, 2024)

#### Domain Challenges : Physical systems that are

(eg., in Computational Fluid Dynamics – CFD, Turbulance, Flows)

- Complex to model/solve analytically
- Computionally expensive to solve numerically

## **Physics and Machine Learning**



Integrating geometric priors in learned representations (Bronstein 2017) Geometric deep learning, **GNN** and neural passing message (Arjona Martínez 2019)

• Differential equations to improve deep learning

Neural differential equations, diffusion models, ...

• Deep learning to solve differential equations

Hypersolvers, hybrid solvers, neural operators, PINNs - Physics-Informed NNetworks, ... (Raissi 2019)

#### $\Rightarrow$ **Promising for engineering**, it allows :

- the integration of analytic knowledge from physical laws governing the engineering systems, to augment statistical knowledge learned from data (eg. by deep learning)
- reducing the high cost of physical simulation, in particular in industry









#### **Scientific Challenges**

- Problems highly-nonlinear, high-dimensional, with complex structures (eg. organized in graphs...)
- Need for adapted NN architectures: GNNs, Deep AEs ..



## Physics-Informed Machine Learning: combining ML and Physics

- Enables prior scientific knowledge based on physics to be taken into account in data-driven machine learning methods  $\rightarrow$ e.g including PINNs - Physics-Informed Neural Nets (Raissi's paper in 2019)
- → Has been successfully and increasingly applied to solve a wide variety of linear and nonlinear problems in physics, covering various fields like mechanics, fluid dynamics, thermodynamics, electromagnetism ... including :



- Solving Navier–Stokes equations coupled with the ۲ corresponding temperature equation for analyzing heat flow convection (NSE+HE). Cai et al, 2021
- Solving incompressible Navier–Stokes equations (NSE). Jin et • al., 2020.
- Solving Euler equations (EE) that model high-speed ۲ aerodynamic flows. Mao et al, 2019
- Solving the nonlinear Shrödinger Equation (SE). •

Raissi, M et al. (2019) Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. Journal of Computational Physics. 378. Online Cuomo, S., et al., (2022). Scientific machine learning through physics-informed neural networks: Where we are and what's next. Journal of Scientific Computing, 92(3), 88. Read Online



## Hybrid ML modeling for solving Partial Differential Equations



A neural framework for solving PDEs, where

- the AI solver is a PINN trained to estimate target function *f*.
- The derivative of **x** is calculated by automatically differentiating the NN's outputs.
- When the differential equation parametrized by  $(\eta)$  is unknown, it can be estimated by solving a loss that optimizes both the functional form of the equation and its fit to observ y.

Eg. Learning Computational Fluid
 Dynamics

- Navier-Stokes Equations: fundamental partial differentials equations (**PDE**) that describe the flow of incompressible fluids.

- Challenge: High-Dimensional non-linear Physical Equations



Credit: Emmanuel Menier, PhD LISN/SystemX

<sup>C.L. M. H. Navier, Memoire sur les Lois du Mouvements des Fluides, Mem. de l'Acad. d. Sci.,6, 398 (1822)
C.G. Stokes, On the Theories of the Internal Friction of Fluids in Motion, Trans. Cambridge Phys. Soc., 8, (1845)</sup> 



**Latent Variable Models**: A family of probabilistic models capable of inferring the intrinsic latent structure (of reduced dimension) of the data

- Auto-Encoders AE (LeCun 1987): The encoder projects the input X (of high-dimension dimension) in a compressed latent representation Z (the code) to reconstruct it using the decoder with outpu X
- → Learning by minimizing the reconstruction error between X and X. The smaller the error, the better the compressed representation Z.
- Variational Auto-encoders VAE (Kingma & Welling 2014) improve the representational capabilities of AEs by regularizing the latent space with a Gaussian prior, coupled with a variational learning
- => can learn complex distributions.
- Deep NNets are excellent candidates







- Nnets are capable to recover highly non-linear relationships in the data
- Adapted architectures that work in a low-dimensional (latent) space



Shallow / Linear

Deep / Non-Linear

Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). Machine learning for fluid mechanics. Annual review of fluid mechanics, 52, 477-508. Read Online 9



## The Research Program IA2: AI and Augmented Engineering



Intelligence artificielle et ingénierie augmentée

#### Artificial Intelligence and Augmented Engineering

- a program with 6
   R&D collaborative
   projects on concrete
   industrial use cases
- Area: Hybrid Al
- 20+ industrial and academic partners
- ~12M€







### HAS Project: Industrial use cases

AIRBUS	AIRBUS: Aircraft aero-acoustic performance	
<b>Air Liquide</b>	Air Liquide: Steam methane reforming	
edf	EDF: Energy modeling of buildings	
	Michelin: 3D rolling tire	
Rte	RTE : Augmented simulator for electrical grid management	
SNCF	SNCF: Forecasting crack propagation in rails	→ B _25 mm

# HSA Project : Simulation/machine learning hybrid modeling

#### Challenges and possible solutions (studied as part of the HSA project):

- Augmenting physical solvers with data-driven models that integrate physics constraints
- Building model architecture adapted to the complex physical structures/systems
- Reducing the simulation cost
- Surrogate models for physical simulation, aiming to Replace physical solvers with
- Deep learning intergrating physical constraints (eg. Deep Graph Nets for PDEs) ➔
- Deal with high-dimensional, non-linear, and complex structurs (e.g reduced modeling, ..)





#### Reduced models and deep learning for PDEs PhD Thesis of E. Menier, 2024 (LISN, Inria/SystemX)

E. Menier et al., 2023. CD-ROM: Complementary Deep-Reduced Order Model. Computer Methods in Applied Mechanics and Engineering 410. Read Online

12

#### High-Dimensional non-linear Physical Equations



#### Deep Graph Neural Networks for Numerical Simulation of PDEs. PhD of W. Liu. 2023 (LISN, Inria/SystemX). Read Online https://www.irt-systemx.fr/projets/HSA/



# Project HSA : simulation and deep learning of graphs

#### Graph Neural Nets for 3D meshes

More suitable, as they operate by construction on graphs

- Generic nature of the learned models
- Transfer learning for improved results
- Prediction can be improved via transfer learning:

from low fidelity (coarse mesh) to high fidelity (finer mesh) models





Modeling of the **airflow** profile around an aircraft wing (Air Foil)

#### Physics: Navier-Stokes equations

Ground Truth

Graph U-Net









Figure 6.3: An example of wheel contact prediction

## **Dynamics: Hybrid ML for HD dynamical physical systems**







High-Dimensional non-linear Dymical Systems:

#### **Goals:**

Recover the dynamics, non-linearity in a high-dimensitonal setting

An Auto-Endore based architecture

The lower-dimensional representation (**z**) is propagated in time using a linear and a nonlinear part based on the Mori-Zwanzig formalism

The decoder D reconstructs the high-dimensional systems.

Menier, E., et al. (2023). Interpretable learning of effective dynamics for multiscale systems. *arXiv preprint arXiv:2309.05*812. <u>Read Online</u> PhD Thesis of E. Menier, 2024 (LISN, Inria/SystemX), 2024



## Validation of hybrid physical-ML systems: LIPS Platform

- How to validate hybrid (with physics) ML approaches ?
  - Several evaluation criteria are required (statistical performance, physical compliance, generalization, etc)
  - Comparison of # ML methods on several specific physical problems => need for a common evaluation framework
- Proposed solution for industry : LIPS "Learning Industrial Physical Simulation" benchmark suite (Result of the project HSA-IA2)
  - 1st framework for the evaluation of physical simulators augmented by machine learning
  - 7 use cases integrated
  - Open-source Framework <u>https://github.com/IRT-SystemX/LIPS Published at NeurIPS2022</u>





LIPS hosts/ed the three following competitions:





System×

Challenges and industrial applications in

- Hybrid AI: How to exploit industry knowledge of physical (scientific) and symbolic (semantic) nature in data-driven learning models.
- Trustworthy AI: Towards integrating AI into critical system engineering
- Generative AI for industry: How to evaluate and advance specialized generative AI models for industrial applications



# Symbolic approaches and Machine Learning (ML and Semantics)

#### → The challenge: **Producing explanations**

There are several ways of producing explanations:

- 1. training the NNet to produce arguments at the same time as prediction
- 2. => or using a **hybrid approach** combining symbolic and neural methods
- Combining learning and knowledge graphs enables business expertise to be integrated and results to be explained.

➔ Neuro-symbolic methods aim at bridging techniques from symbolic AI and deep learning:

- integrating a symbolic paradigm into a neural network
- E.g The ML model uses explicit symbolic knowledge, in the form of logic rules/ontologies, to specify desired properties for the NNet.
  - Formal rules for explainability (Audemard, 2023)

• Improving Neural-Based Classification with Logical Background Knowledge (Ott 2023) (Battagla 2018)



## Integrating Expert/Business <u>semantics</u> in ML (Project SMD)

#### The Neuro-symbolic engineering pipeline:

https://www.irt-systemx.fr/projets/SMD/



Framework:

Building ML-based on knowledge graphs from expert/business language data

**Approach** (The Neuro-symbolic engineering pipeline)

- Information extraction from heterogeneous structured and semi-structured text corpora...:
- NLP approach; Semantic annotation
- Taking into account domain rules/constraints in the numerical-AI based decision

Intelligence artificielle et ingénierie augmentée



AIRBUS

## Integrating Expert/Business semantics in ML (Project SMD)



Logic rules, to specify desired properties for the Nnet:



MODEL

- => Insertion of logical rules at the inference step :
- Recalculate  $P(y|X, \alpha)$  in lieu of the learnt P(y|X)
- $\alpha$  is a rule encoding the validity of the prediction  $\hat{y}$

string



string

Ontologies and (machine/transfer) learning for multimedia document analysis. PhD thesis of A. Ledaguenel (in progress, MICS/SystemX)

### Some use cases



**Use case** : Detection and characterization of atypical scenes in surveillance videos for crisis management









ShanghaiTech data (437 videos on 13 different locations, containing 130 abnormal events)

- identifying anomalies in video streams
- scene analysis and construction of contextualized graphs









Road accident

UCF-crime data (128 hours videos contain realistic anomalies including Abuse, Arrest, Arson, Assault, etc.)



### Challenges and industrial applications in

- Hybrid AI: How to exploit industry knowledge of physical (scientific) and symbolic (semantic) nature in data-driven learning models.
- Trustworthy AI: Towards integrating AI into critical system engineering
- Generative AI for industry: How to evaluate and advance specialized generative AI models for industrial applications



Confiance.Al programme: A French unique community to design and industrialise trustworthy Al-based critical systems





### Some Use Cases

In distribution







Valeo



Theme	Primary	Secondary
2D Vision	Scene understanding (Valeo) In autonomous driving	Aerial pictures (Thales LAS)
Visual inspection	Welding inspection (Renault)	Industrial control (Safran)
Time series prediction	<b>Demand forecasting</b> (Air Liquide) Eg. oxygen	
Time series anomaly detection	Plant efficiency monitoring (Air Liquide)	Virtual sensor (Airbus Helicopter)
Tabular data	ACAS XU (Airbus)	
NLP	Opinion mining (Renault)	
Hybrid ML Symbolic	Time dependent planning (Safran)	
Visual similarity	Re-identification (ATOS)	



Out distribution





### **AIRBUS**

Credit : Confiance.ai



## Scientific challenges of confiance.ai

- **confiance.ai:** methods and tools for trusted AI.
- High expectations for industry
- **3 groups** of **scientific challenges** to cover all aspects of trust
  - 1) Trust and learning data
  - 2) Trust and human interaction
  - 3) Trust and AI-based system engineering
- Organization in **7 projects**:
  - > EC1: Integration & **use cases**
  - > EC2: Process, **methodology** and guidelines
  - > EC3: Characterization & qualification of trustworthy AI
  - EC4: **Design** for Trustworthy AI
  - > EC5: Data, information and knowledge **engineering** for trusted AI
  - > EC6: IVV&Q strategy toward homologation / certification
  - > EC7: Target Embedded AI

Confion



- confiance.ai: methods and tools for trusted AI → High expectations for industry
- many scientific challenges (as part of the confiance.ai doctoral programme)
  - Data coverage and operational domain design ODD in Computer Vision (PhD Theis of Adrien Le Coz)
  - **Robustness** to 'patch' adversarial attacks in Computer Vision (PhD Thesis of Paul La Barbarie)
  - Robust **Reinforcement** learning (PhD Thesis of Lucas Schott)
  - Statistical control of Fairness/Bias in Machine Learning (Stat) (PhD Thesis of Gayane Taturyan)
  - Towards hardware **deployment** using BO & Neural Architecture Search) (PhD Theis of H Ouertatani)
  - Monitoring of (deployed) AI algorithms using conformal prediction (PhD Thesis of Abdelmouaiz Tebjou)



### Confiance.ai : Key figures







#### Challenges and industrial applications in

- Hybrid AI: How to exploit industry knowledge of physical (scientific) and symbolic (semantic) nature in data-driven learning models.
- Trustworthy AI: Towards integrating AI into critical system engineering
- Generative AI for industry: How to evaluate and advance specialized generative AI models for industrial applications



- **\*GANs** (Goodfellow 2014): A **generator G** and a **discriminator D** subjected to two contradictory training (ie. *adversarial* aspect )
- → impressive results in images,...
- → IRT work advancing the state of the art on the subject: in particular, work on EPI, RTI Confiance.AI/EC3,4,5 projects.

\***Diffusion models** (Ho 2020; Yang 2023; Rombach 2022): an unintuitive principle at the basis : progressively **destructure** the input

until it is completely degraded, then **reconstruct** it by reversing the process.

- Excellent performance in image synthesis, despite costly training
- ➔ Research work atIRT (Confiance.Al, EC4/Explo; e.g. NeurIPS'23)



\***Transformers** (Vaswani 2017): have revolutionized generative AI, particularly in NLP.

- ➔ highly parallelizable (unlike sequential architectures like RNNs)
- → Use an AE and relies on an attention mechanism to integrate global I/O and context dependencies into a variable-length sequence.
- → High NLP capabilities {eg. ChatGPT-3.5: responses of up to ~ 3000 words}.
- NLP models and transformers are studied in R&D at IRT
- (eg. Confiance.Al, SMD)



Yang Song Blog

**OpenAl** 



- Main scientific challenge in GenAl in industry (Hybridization, Frugality, Multimodalities, Evaluation/Benchmarking, ..)
  - the **frugality** of foundation models (related to data, model, and learning),
  - operating on different multi-modalities (beyond text : time-series, diagrams, images..),
  - their hybridization to integrate knowledge (expert and or scientific knowledge),
  - their specialization (eg. fine-tuning, RAG) to different use cases,
  - and their evaluation to guarantee industrial use,
- Preparation to the implementation of the AI act
- Hybrid AI (GenAI, Physics-Simulation, Data augmentation, Uncertainty Quantification, ..)





- Hybrid AI, with its inclusive approach to human knowledge, overcomes the limitations of "classic" AI (based exclusively on data) : The
  - accuracy of physical simulations can be improved by hybrid modeling that takes advantage of data.
  - joint use of data and scientific laws reduces the complexity and cost of physical simulations.
  - joint use of data & knowledge graphs enables business expertise to be integrated and results to be explained.
- Approaches that hybridize data and knowledge models (physical/semantic) have emerged fairly recently, and have not yet reached maturity ==> R&D efforts are needed to bring the subject to maturity in engineering/industry.
- An avenue in AI to the preparation of the implementation of the AI act for the entreprise, is Trusted AI.
   Need for the Human in the loop
- **Generative AI** is beginning to make slight progress in industry...



### Some references

#### **General hybrid**

- Wang & al. (2023). Scientific discovery in the age of artificial intelligence. Nature, 620
- L. von Rueden et al. (2023) Informed Machine Learning A Taxonomy and Survey of Integrating Prior Knowledge into Learning Systems, IEEE Transactions on Knowledge and Data Engineering, vol. 35(1)

#### **Physics/Simulation and ML**

- Bronstein, M. M. & al. (2017) Geometric deep learning: going beyond euclidean data. IEEE Signal Process Mag. 34
- Arjona Martínez, J., et al. (2019) Pileup mitigation at the Large Hadron Collider with graph neural networks. Eur. Phys. J. Plus 134, 333
- Monti, F., et al. (2016): Geometric deep learning on graphs and manifolds using mixture model CNNs, CVPR.
- Davies, A. et al. (2021) Advancing mathematics by guiding human intuition with Al. Nature 600
- Raissi, M et al. (2019) Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. Journal of Computational Physics. 378
- Liu et al. (2020), Energy-based Out-of-distribution Detection, NeurIPS
- E. Menier, et al. (2023) CD-ROM: Complemented Deep Reduced order model. Computer Methods in Applied Mechanics and Engineering. Vol 410.

#### Smentics/Symbolics and ML

- T. Laugel et al. (2019) The Dangers of Post-hoc Interpretability: Unjustified Counterfactual Explanations. Int. Joint Conference on Artificial Intelligence
- A. Pirovano, et al. (2021) Computer aided diagnosis tool for cervical cancer screening with weakly supervised localization and detection of abnormalities using adaptable and explainable classifier. Medical Image Analysis, 73
- M. Charachon et al. (2020), Combining similarity and adversarial learning to generate visual explanation : Application to medical image classification. ICPR
- K. Marino et al. (2017) The more you know: Using knowledge graphs for image classification, Conf. Computer Vision and Pattern Recognition (CVPR)
- P. W. Battaglia et al. (2018) Relational inductive biases, deep learning, and graph networks," arxiv:1806.01261
- E. Marconato et al. (2022) GlanceNets: Interpretable, Leak-proof Concept-based Models. Neurips
- G. Audemard, et al. (2023) Deriving Explanations for Decision Trees: The Impact of Domain Theories. hal-04176274
- J. Ott et al. (2023) How to Think About Benchmarking Neurosymbolic AI? 17th Intern. Workshop on Neural-Symbolic Learning and Reasoning (NeSy)



### Some references

- Marot, A., et al. (2018). Guided machine learning for power grid segmentation. In 2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe) (pp. 1-6).
- Merino-Martínez et al. CEAS Aeronautical Journal (2019).
- Raissi, M et al. (2019) Physics-Informed Neural Networks: A Deep Learning Framework for Solving Forward and Inverse Problems Involving Nonlinear Partial Differential Equations. Journal of Computational Physics. 378. Online
- Cuomo, S., et al., (2022). Scientific machine learning through physics—informed neural networks: Where we are and what's next. *Journal of Scientific Computing*, 92(3), 88. <u>Read Online</u>
- C.L. M. H. Navier, Memoire sur les Lois du Mouvements des Fluides, Mem. de l'Acad. d. Sci., 6, 398 (1822)
- C.G. Stokes, On the Theories of the Internal Friction of Fluids in Motion, Trans. Cambridge Phys. Soc., 8, (1845)
- Wang & al. (2023). Scientific discovery in the age of artificial intelligence. Nature, 620. Read Online
- Brunton, S. L., Noack, B. R., & Koumoutsakos, P. (2020). Machine learning for fluid mechanics. Annual review of fluid mechanics, 52, 477-508. Read Online
- Donon, B., et al. (2020). Deep statistical solvers. Advances in Neural Information Processing Systems, 33, 7910-7921. Read Online
- Donon, B. (2022). Deep statistical solvers & power systems applications (Doctoral dissertation, Université Paris-Saclay). Read Online
- E. Menier et al., 2023. CD-ROM: Complementary Deep-Reduced Order Model. Computer Methods in Applied Mechanics and Engineering 410. Read Online
- Deep Graph Neural Networks for Numerical Simulation of PDEs. PhD of W. Liu. 2023 (LISN, Inria/SystemX). Read Online
- Ho et al. 2022. Denoising diffusion probabilistic models. Adv. NeurIPS., 33 (2020), pp. 6840-6851. Read Online
- D. Shu et al., 2023. A physics-informed diffusion model for high-fidelity flow field reconstruction, Journal of Computational Physics, 478, 2023. Read Online
- Liu et al. (2020), Energy-based Out-of-distribution Detection, NeurIPS. Read Online
- D. Shu et al., 2023. A physics-informed diffusion model for high-fidelity flow field reconstruction, Journal of Computational Physics, 478, 2023. Read Online



Accélérateur de la transformation numérique



# THANK YOU FOR YOUR ATTENTION

