On some modern statistical and machine learning approaches for industrial applications

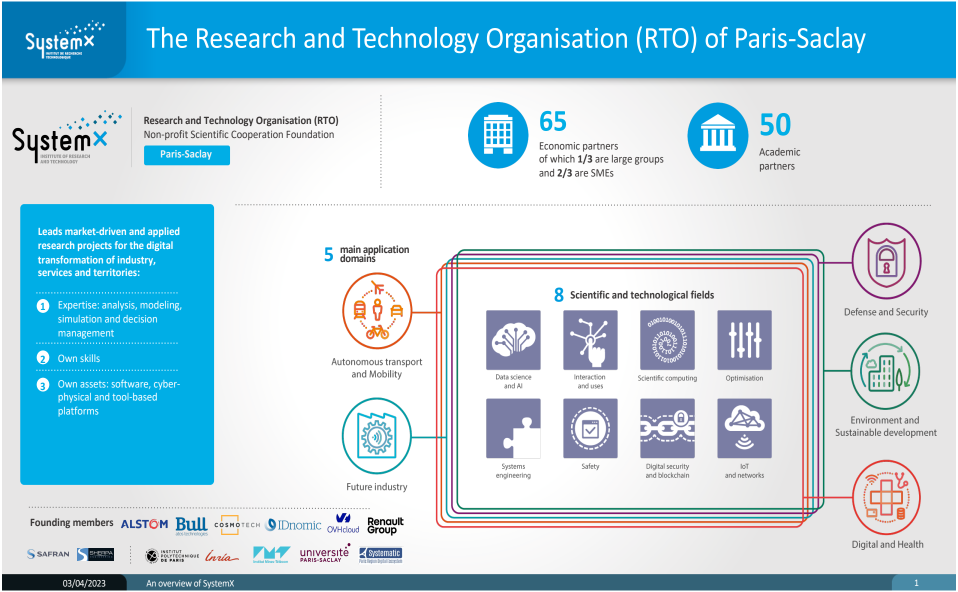
Faïcel Chamroukhi

SystemX
INSTITUT DE RECHERCHE TECHNOLOGIQUE

Talk at:

4ème CISEM
Colloque International Statistique et Économétrie
Mahdia, 17 - 19 Mai 2024

IRT SystemX



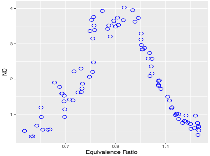
The Research and Technology Organisation (RTO) of Paris-Saclay
Research and Technology Organisation (RTO)
Non-profit Scientific Cooperation Foundation
Paris-Saclay
65
Economic partners
of which 1/3 are large groups
and 2/3 are SMEs
50
Academic
partners
Leads market-driven and applied research projects for the digital transformation of industry, services and territories:
1 Expertise: analysis, modeling, simulation and decision management
2 Own skills
3 Own assets: software, cyber-physical and tool-based platforms
5 main application domains
Autonomous transport and Mobility
Future industry
8 Scientific and technological fields
Data science and AI
Interaction and uses
Scientific computing
Optimisation
Systems engineering
Safety
Digital security and blockchain
IoT and networks
Defense and Security
Environment and Sustainable development
Digital and Health
Founding members
03/04/2023
An overview of SystemX
1

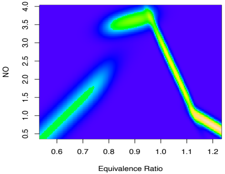
Outline

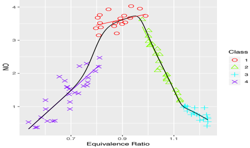
- Part 1 : mixtures-of-experts for heterogenous and high-dimensional data
- Part 2 : challenges and industrial applications in Hybrid & Trustworthy AI

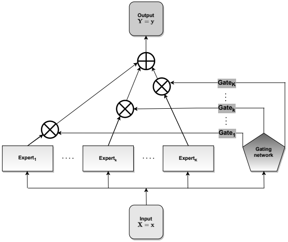
Mixtures-of-Experts to model heterogeneous data

Mixtures-of-Experts as good candidates to model a response Y given predictor.s X governed by a hidden structure accounting for heterogeneity









Schematic diagram of the neural network architecture of a K-component MoE model.

first studied as neural networks (NNs) by Jacobs, Jordan, Nowlan, and Hinton (1991)
HD Nguyen and F Chamroukhi. Practical and theoretical aspects of mixture-of-experts modeling : An overview. WIRES : Data Mining and Knowledge Discovery Wiley Periodicals, Inc. 2018

Density approximation in Unsupervised Learning

- **Data** : observations $\{\boldsymbol{x}_i\}$ from $\boldsymbol{X} \in \mathbb{X} \subset \mathbb{R}^d$ of density (multimodal) $f \in \mathcal{F}$
- **Objective** : approximate the density f (and represent the data, e.g. *clustering*)
- **Solution :** Approximate f within the class $\mathcal{H}^\varphi = \bigcup_{K\in\mathbb{N}^\star} \mathcal{H}_K^\varphi$ of finite location-scale mixture h_K^φ (of K-components) of density φ (e.g., Gaussian), where

$$\mathcal{H}_K^\varphi = \left\{ \boxed{h_K^\varphi(\boldsymbol{x}) := \sum_{k=1}^{K} \pi_k \frac{1}{\sigma_k^d} \varphi\left(\frac{\boldsymbol{x} - \boldsymbol{\mu}_k}{\sigma_k}\right)}, \boldsymbol{\mu}_k \in \mathbb{R}^d, \sigma_k \in \mathbb{R}_+, \pi_k > 0\, \forall k \in [K], \sum_{k=1}^{K} \pi_k = 1 \right\}$$

Theorem : Universal approximation of finite location-scale mixtures

(a) Given any p.d.f $f, \varphi \in \mathcal{C}$ and a compact set $\mathbb{X} \subset \mathbb{R}^d$, there exists a sequence $(h_K^\varphi) \subset \mathcal{H}^\varphi$, such that $\lim_{K\to\infty} \sup_{\boldsymbol{x}\in\mathcal{X}} |f(\boldsymbol{x}) - h_K^\varphi(\boldsymbol{x})| = 0$.

(b) For $p \in [1, \infty)$, if $f \in \mathcal{L}_p$ (Lebesgue p.d.f) and $\varphi \in \mathcal{L}_\infty$ (essentially bounded p.d.f), there exists a sequence $(h_K^\varphi) \subset \mathcal{H}^\varphi$, such that $\lim_{K\to\infty} \|f - h_K^\varphi\|_{\mathcal{L}_p} = 0$.

Nguyen, T., Chamroukhi, F., Nguyen, H. D., & McLachlan, G. J. (2023). Approximation of probability density functions via location-scale finite mixtures in Lebesgue spaces. Communications in Statistics - Theory and Methods, 52(14), 5048–5059. https://arxiv.org/pdf/2008.09787

Modeling with mixtures-of-experts (ME)

- **Context :** n observations $\{\boldsymbol{x}_i, \boldsymbol{y}_i\}$ from a pair $(\boldsymbol{X}, \boldsymbol{Y}) \in \mathbb{X} \times \mathbb{Y}$ with unknown conditional p.d.f $f \in \mathcal{F} = \left\{f : \mathbb{X} \times \mathbb{Y} \to \mathbb{R}_+ \mid \int_{\mathbb{Y}} f(\boldsymbol{y}|\boldsymbol{x}) \, \mathrm{d}\lambda(\boldsymbol{y}) = 1, \forall \boldsymbol{x} \in \mathbb{X}\right\}$
- **High-dimensional setting** : $\mathbb{X} \subseteq \mathbb{R}^d$, $\mathbb{Y} \subseteq \mathbb{R}^q$, with $d, q \gg n$ and **heterogeneous**.
- **Objectives :** Regression ; Clustering ; Model selection
- **Solution :** Approximate f within the class of **mixtures-of-experts** :

Let φ be a p.d.f (compactly supported on $\mathbb{Y} \subseteq \mathbb{R}^q$), we define the functional classes :

- Location-scale family : $\mathcal{E}_\varphi = \left\{\phi_q(\boldsymbol{y}; \boldsymbol{\mu}, \sigma) := \frac{1}{\sigma^q} \varphi\left(\frac{\boldsymbol{y} - \boldsymbol{\mu}}{\sigma}\right); \boldsymbol{\mu} \in \mathbb{Y}, \sigma \in \mathbb{R}_+\right\}$.
- Mixture of location-scale experts with softmax activation network : SGaME :

$$\mathcal{H}_S^\varphi = \left\{\boxed{h_K^\varphi(\boldsymbol{y}|\boldsymbol{x}) := \sum_{k=1}^{K} g_k(\boldsymbol{x}; \boldsymbol{\gamma}) \phi_q(\boldsymbol{y}; \boldsymbol{\mu}_k, \sigma_k)}; \quad \phi_q \in \mathcal{E}_\varphi \cap \mathcal{L}_\infty, g_k(\cdot; \boldsymbol{\gamma}) \in \{\text{softmax}\}\right\}$$

Theorem : Approximation capabilities of isotropic mixtures-of-experts SGaME

(a) For $p \in [1, \infty)$, $f \in \mathcal{F}_p \cap \mathcal{C}$, $\varphi \in \mathcal{F} \cap \mathcal{C}$, $\mathbb{X} = [0,1]^d$, there exists a sequence $(h_K^\varphi) \subset \mathcal{H}_S^\varphi$ such that $\lim_{K \to \infty} \left\| f - h_K^\varphi \right\|_{\mathcal{L}_p} = 0$.

(b) For $f \in \mathcal{F} \cap \mathcal{C}$, if $\varphi \in \mathcal{F} \cap \mathcal{C}$, $d = 1$, there exists a sequence $(h_K^\varphi) \subset \mathcal{H}_S^\varphi$ such that $\lim_{K \to \infty} h_K^\varphi = f$ almost uniformly.

Nguyen, H.D., Nguyen, T., Chamroukhi, F., McLachlan G. J. Approximations of conditional probability density functions in Lebesgue spaces via mixture of experts models. Journal of Statistical Distributions and Applications. 8, 13 (2021). https://doi.org/10.1186/s40488-021-00125-0

Learning with mixtures-of-experts models

Learning via the EM algorithm : $\boldsymbol{\theta}^{new} \in \arg\max_{\boldsymbol{\theta}\in\Omega} \mathbb{E}[\ln L_c(\boldsymbol{\theta})|\mathcal{D}, \boldsymbol{\theta}^{old}]$

SaMUraiS : open source software for statistical time-series analysis

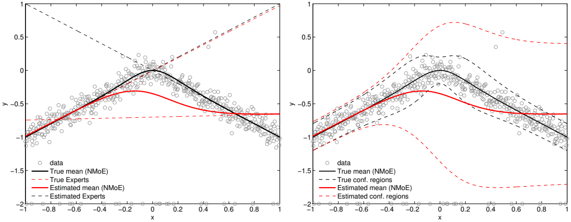
SaMUraiS : **S**t**A**tistical **M**odels for the **U**nsupe**R**vised segment**A**t**I**on of time-**S**eries

▸ Github ▸ CRAN ▸ Matlab software

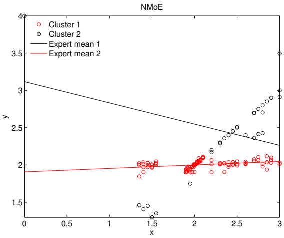
Available algorithms and Packages

RHLP : Regression with Hidden Logistic Process	▸ R software	▸ Matlab software
HMMR : Hidden Markov Model Regression	▸ R software	▸ Matlab software
PWR : Piece-Wise Regression	▸ R software	▸ Matlab software
MRHLP : Multivariate RHLP	▸ R software	▸ Matlab software
MHMMR : Multivariate HMMR	▸ R software	▸ Matlab software
MPWR : Multivariate PWR	▸ R software	▸ Matlab software

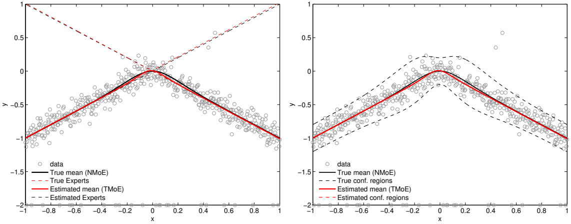
Include estimation, segmentation, approximation, model selection, and sampling



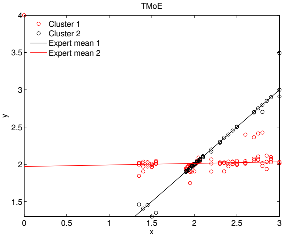
$n = 500$ observations with 5% of outliers $(x; y = -2)$: Normal fit



Tone data with 10 outliers $(0, 4)$: Normal fit



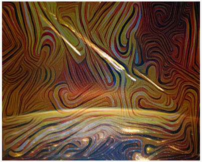
$n = 500$ observations with 5% of outliers $(x; y = -2)$: **Robust** fit



Tone data with 10 outliers $(0, 4)$: **Robust** fit

Robust learning with mixtures-of-experts models

MEteorits : open-source soft. Robust learning with mixtures-of-experts models



MEteorits : **M**ixtures-of-**E**xper**T**s mod**E**ling for c**O**mplex and non-no**R**mal d**I**s**T**ributionS

▸ Github ▸ CRAN ▸ Matlab software

Available algorithms and Packages

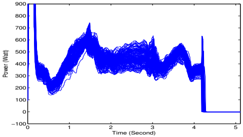
`NMoE` : Normal Mixture-of-Experts	▸ R software	▸ Matlab software
`SNMoE` : Skew-Normal Mixture-of-Experts	▸ R software	▸ Matlab software
`tMoE` : Robust MoE using the t-distribution	▸ R software	▸ Matlab software
`StMoE` : Skew-t Mixture-of-Experts	▸ R software	▸ Matlab software

- Meteorits include sampling, fitting, prediction, clustering with each MoE model
- Non-normal mixtures (and MoE) is a very recent topic in the field

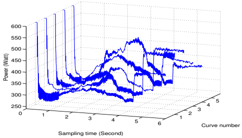
Some real-world data

- Heterogenous, Multimodal, High-Dimensional, Unlabeled, Possibly Massive ...
- Need for adapted analysis tools

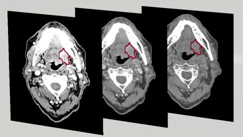
Transport : Railway switch curves diagnostic



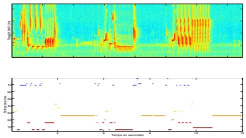
Predictive Maintenance

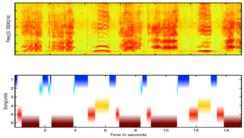


Health : Medical images

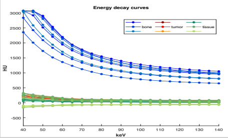


Acoustics : scene listening (marine, terrestrial)

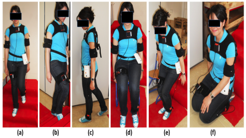




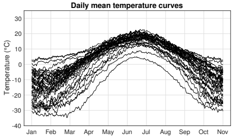
Dual-energy computed tomography

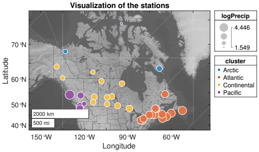


Health & Well Being : Activity recog.



Climate/Environment : meteorological data



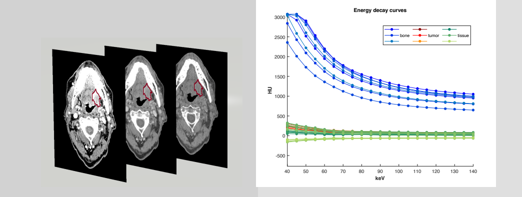


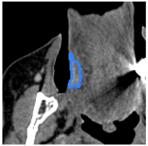
Dual-energy computed tomography (DECT) image Clustering

- Learning from Multimodal information in Healthcare/Radiology
- Cancer detection in Radiology : DECT clustering [Diagnostics (AI in medicine), 2022]

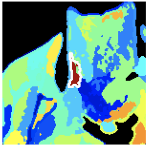
  Spatial mixture of functional regressions for dual-energy CT images

  $m(\boldsymbol{y}|\boldsymbol{x}, \boldsymbol{v}; \boldsymbol{\theta}) = \sum_{k=1}^{K} \alpha_k(\boldsymbol{v}; \boldsymbol{\alpha}) f_k(\boldsymbol{y}|\boldsymbol{x}; \boldsymbol{\theta}_k)$ where $\alpha_k(\boldsymbol{v}; \boldsymbol{\alpha}) = \frac{w_k \phi_3(\boldsymbol{v}; \boldsymbol{\mu}_k, \mathbf{R}_k)}{\sum_{\ell=1}^{K} w_\ell \phi_3(\boldsymbol{v}; \boldsymbol{\mu}_\ell, \mathbf{R}_\ell)}$



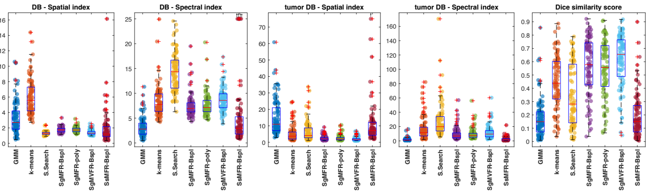


(a) Original slice



(b) Dice=0.84, DB=1.64/6.92

DECT multimodal Data : 3D voxels & energy levels Expert Annotation Automatic Annotation

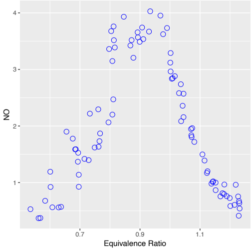


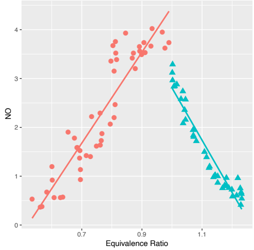
▸ Codes available on Github

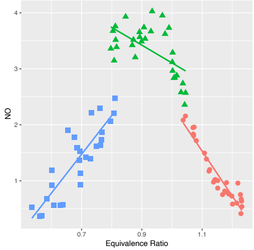
Model estimation and selection in MoE

(a) Raw Ethanol data set

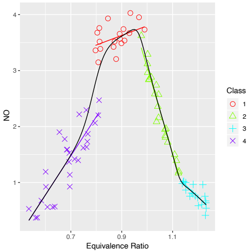
Collection of MoE models with linear mean functions characterized by 2-5 clusters

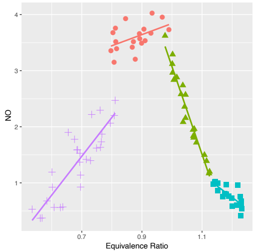


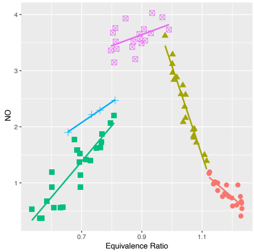




(b) Best data-driven MoE model







Learning with high-dimensional predictors

Questioning : Prediction (non-linear regr., classification) & clustering in presence of

[1.] **High-dimensional** predictors : $\boldsymbol{X}_i \in \mathbb{R}^p$ with $p \gg n$

[2.] **Functional** predictors : $X_i(t)$, $t \in \mathcal{T} \subseteq \mathbb{R}$ {eg. continuously recorded variables}

$\hookrightarrow$ Look for parsimonious and interpretable methods

[1.] HDME : High-Dimensional Mixtures-of-Experts

- Learning : PMLE $\widehat{\boldsymbol{\theta}}_{\mathrm{n}} \in \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{n} \log h_K^{\varphi}(\boldsymbol{y}_i|\boldsymbol{x}_i;\boldsymbol{\theta}) - \mathrm{pen}(\boldsymbol{\theta})$
- $\hookrightarrow$ Lasso penalty : $\mathrm{Pen}_\lambda(\boldsymbol{\theta}) = \underbrace{\sum_{k=1}^{K} \lambda_k \|\boldsymbol{\beta}_k\|_1}_{\text{Experts Net.}} + \underbrace{\sum_{k=1}^{K-1} \gamma_k \|\boldsymbol{w}_k\|_1}_{\text{Gating Net.}}$

$\hookrightarrow$ encourages sparse solutions & performs estimation and feature selection

$\hookrightarrow$ computationally attractive (Avoid matrix inversion ; univariate updates)

▸ Software Toolbox HDME on Github (GaussRMoE, LogisticRMoE, PoissonRMoE)

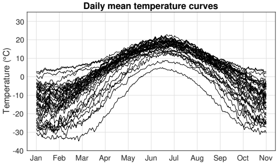
$\hookrightarrow$ A non-asymptotic result. If $\mathrm{pen}(\mathbf{m})$ is well chosen, then our PMLE behaves in a comparable manner compared to **the best (oracle) model** $\mathcal{H}_{\mathbf{m}^\star}$ in the collection

Nguyen TT, Nguyen HD, Chamroukhi F and Forbes F. A non-asymptotic approach for model selection via penalization in high-dimensional mixture of experts models. Electronic Journal of Statistics. 2022

C & &Huynh. Regularized Maximum Likelihood Estimation and Feature Selection in Mixtures-of-Experts Models. Journal de la Société Francaise de Statistique, Vol. 160(1), pp :57–85, 2019

Huynh & C. Estimation and Feature Selection in Mixtures of Generalized Linear Experts Models. arXiv :1810.12161

[2.] Learning with functional predictors



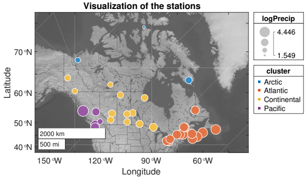


FIGURE – $n = 35$ daily mean temperature measurement curves (X_i's) in different stations (Left) and the log of precipitation values (Y_i's) visualized with the climate regions (Z_i's) (Right).

- Relate functional predictors $\{X(t) \in \mathbb{R}; t \in \mathcal{T} \subset \mathbb{R}\}$ to a scalar response $Y \in \mathcal{Y} \subset \mathbb{R}$
- Regression and classification of heterogeneous responses given functional predictors

  (1) generative functional modeling, sparsity and feature selection (high-dimension)

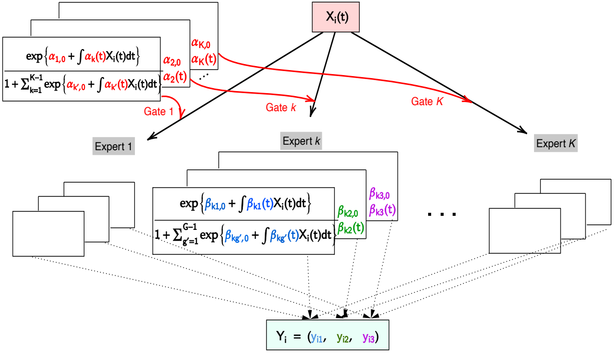
  (2) User guideline : keep an interpretable fit

[2.] Functional Mixtures-of-Experts (and Different Learning strategies, in particular)

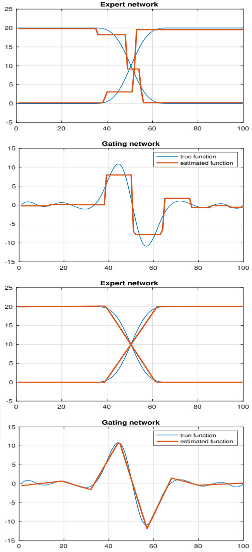
- $Y_i = \beta_{z_i,0} + \int_{\mathcal{T}} X_i(t)\beta_{z_i}(t)dt + \varepsilon_i$ avec $h_z(X_i(.)) = \alpha_{z_i,0} + \int_{\mathcal{T}} X_i(t)\alpha_{z_i}(t)dt$
- Lasso-type Regularized MLE w.r.t the derivatives of the $\alpha(\cdot)$ and $\beta(\cdot)$ functions

C, Pham, Hoang, McLachlan. Functional Mixtures-of-Experts. *Statistics and Computing* ., Vol. 34 (98), 2024 [open access]
PhD Thien Pham, 2022

Mixture-of-Experts Architecture



Interpretable fits



$Y_i = \beta_{z_i,0} + \int_{\mathcal{T}} X_i(t)\beta_{z_i}(t)dt + \varepsilon_i$ with $h_z(X_i) = \alpha_{z_i,0} + \int_{\mathcal{T}} X_i(t)\alpha_{z_i}(t)dt$
l_1-Regularized MLE w.r.t the derivatives of the $\alpha(\cdot)$ and $\beta(\cdot)$ functions

↪ produces a meaningful sparse estimates for $\beta_{z_i}(t)$ curves :

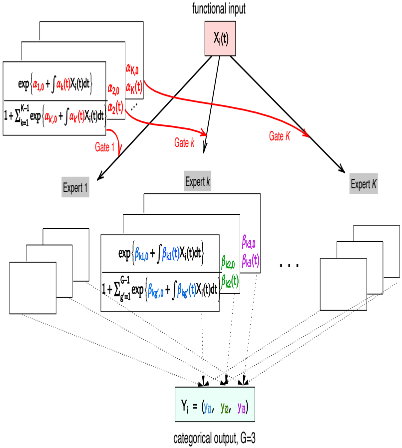
$\beta^{(0)}_{z_i}(t) = 0$ implies that $X(t)$ has no effect on Y at t

$\beta^{(1)}_{z_i}(t) = 0$ means that $\beta_{z_i}(t)$ is constant at t,

$\beta^{(0)}_{z_i}(t) = 1$ shows that $\beta_{z_i}(t)$ is a linear function of t, etc.

C, Pham, Hoang, McLachlan. Functional Mixtures-of-Experts. *Statistics and Computing* ., Vol. 34 (98), 2024 [open access]

Mixture-of-Experts Architecture



produces a meaningful sparse estimates for $\beta_{z_i}(t)$ curves :

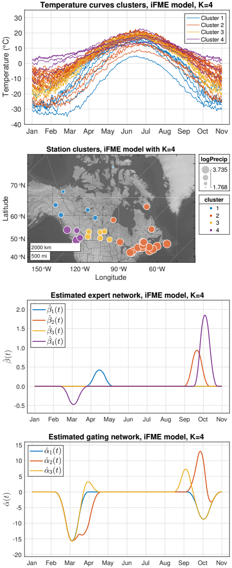
$\beta^{(0)}_{z_i}(t) = 0$ implies that $X(t)$ has no effect on Y at t

$\beta^{(1)}_{z_i}(t) = 0$ means that $\beta_{z_i}(t)$ is constant at t,

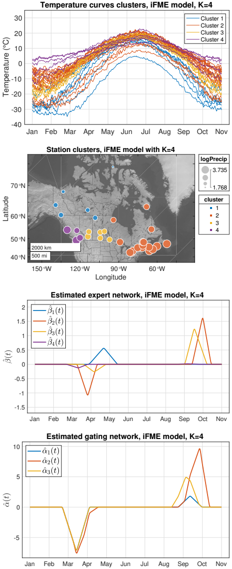
$\beta^{(0)}_{z_i}(t) = 1$ shows that $\beta_{z_i}(t)$ is a linear function of t,

etc.

OUR regularization

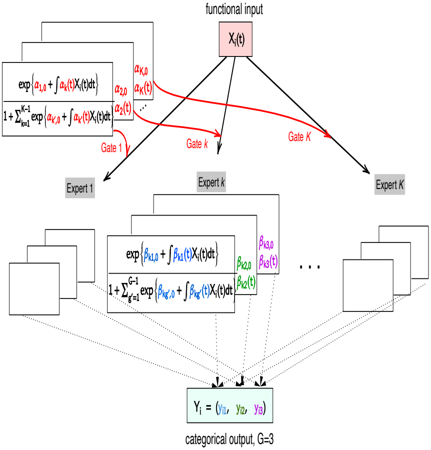


OUR regularization

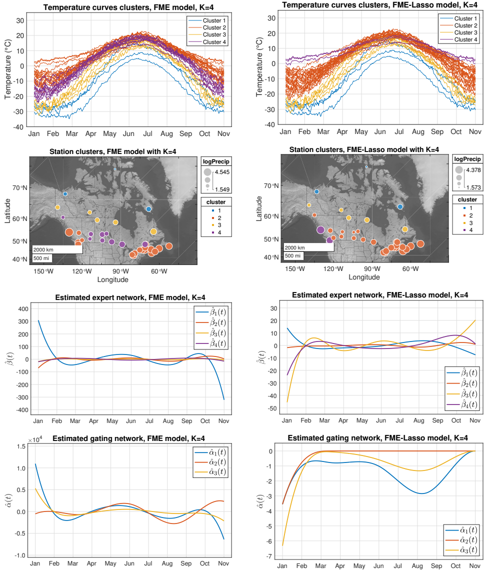


C, Pham, Hoang, McLachlan. Functional Mixtures-of-Experts. *Statistics and Computing* ., Vol. 34 (98), 2024 [open access]

Mixture-of-Experts Architecture



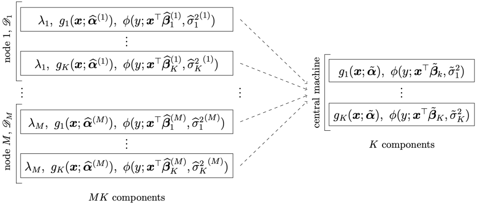
No regularization | LASSO regularization



C, Pham, Hoang, McLachlan. Functional Mixtures-of-Experts. *Statistics and Computing* ., Vol. 34 (98), 2024 [open access]

Aggregating distributed mixtures-of-experts models (MoE)

- collaborative MoE for distributed (eg. large-scale data) or federated learning



- Local estimators : $\hat{f}_m = f(\cdot|\mathbf{x},\widehat{\boldsymbol{\theta}}_m) = \sum_{k=1}^K g_k(\mathbf{x},\widehat{\boldsymbol{\alpha}}^{(m)})\phi(\cdot;\mathbf{x}^\top\widehat{\boldsymbol{\beta}}_k^{(m)},\widehat{\sigma}_k^{2\,(m)})$,
- weighted average : $\bar{f} = f(y|\mathbf{x};\bar{\boldsymbol{\theta}}) = \sum_{m=1}^M \lambda_m \hat{f}_m$ where $\lambda_m = \frac{N_m}{N}$ the sample proportion. $\bar{f}$ is good but relates MK components so not our direct target.

$\hookrightarrow$ Reduced estimator : $\bar{f}^R = \underset{h_K\in\mathcal{M}_K}{\arg\inf}\ \rho\left(h_K,\sum_{m=1}^M \lambda_m \hat{f}_m\right)$: we seek for a K-component ME h that is closest to the MK-component ME $\bar{f} = \sum_{m=1}^M \lambda_m \hat{f}_m$ w.r.t a transportation divergence $\rho(\cdot,\cdot)$, e.g. KL.

C and Phamn T. Distributed Learning of Mixtures of Experts. *arxiv 2312.09877*, 2024 [under revision at IEEE TNNLS]
{PhD, Pham. 2022}

Source codes publicly available on Github.

Numerical results in Distributed clustering and Prediction

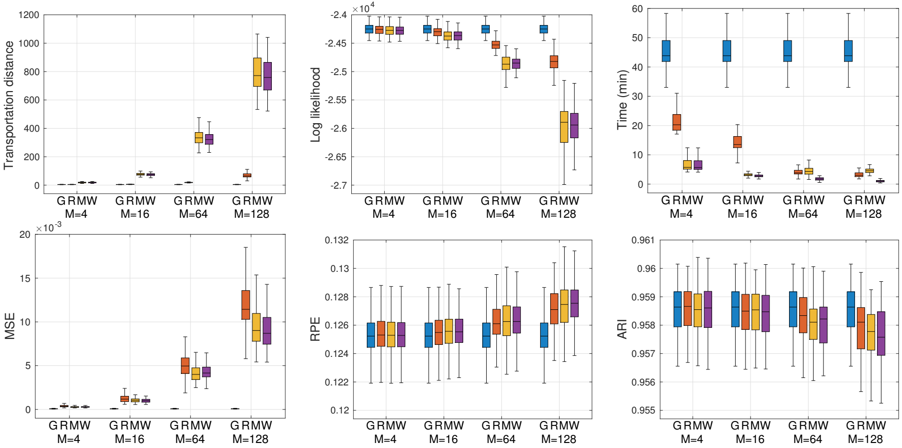


FIGURE – Performance of the Global ME (G), Reduction (R), Middle (M) and Weighted average (W) estimator for sample size $N = 10^6$ and M machines.

C and Phamn T. Distributed Learning of Mixtures of Experts. *arxiv 2312.09877*, 2024 [under revision at IEEE TNNLS] {PhD, Pham. 2022}

▸ Source codes publicly available on Github.

Systemx
INSTITUTE OF RESEARCH AND TECHNOLOGY

Boosting
digital transformation

Challenges and industrial applications in Hybrid AI Trustworthy AI

Faïcel Chamroukhi

4ème CISEM
Colloque International Statistique et Économétrie
Mahdia, 17 - 19 Mai 2024

FRANCE 2030

www.irt-systemx.fr/en

Outline

Challenges and industrial applications in

- Hybrid AI: exploiter les connaissances métiers de natures physique et symbolique dans les modèles d'apprentissage

- Trustworthy AI: concevoir et industrialiser des systèmes à base d'intelligence artificielle de confiance

The Research Program IA2: AI and Augmented Engineering

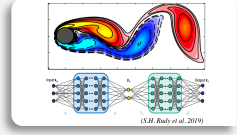
Intelligence artificielle et ingénierie augmentée

Artificial Intelligence
an Augmented Engineering

- a program with **6 R&D collaborative projects** based on concrete **industrial use cases**
- Area: Hybrid AI
- 20+ industrial and academic partners

Advance project
Thesis / Postdocs / Shared work

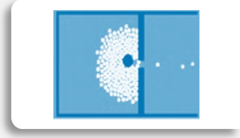
HSA: Simulation/machine learning hybrid modeling 01
How industrial solvers and learning models can enrich each other ?



AFS: Agility and fidelity of simulations 02
How to imporve agility and fidelity of simulation in complex systems design?

S2I: Industrial infrastructure supervision 03
How to improve decision-making on distubuted industrial systems via machine learning techniques ?

SAA: Augmented multi-agent simulation 04
How can multi-agent models benefit from real data and bring out atypical situations?



SMD: Business Semantics for Multi-source Data Mining 05
How to link heterogeneous data with established practical knowledge?

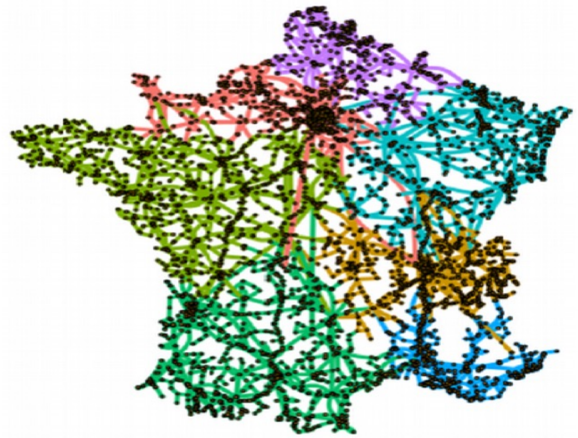
CAB: Cockpit and Bidirectional Assistant 06
How to develop a virtual assistant that learns from expert and learns to the expert

Diagram/pictures credit to [IA2 Program]

Some physical problems in Industry

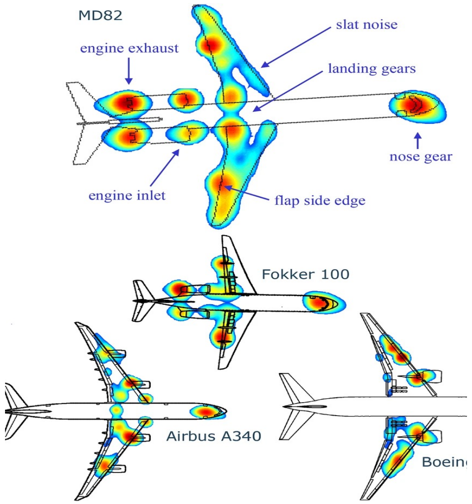
- Related to the desing and supervision of complex (physical) systems
- Covering various fields in physics (mechanics, fluid dynamics, aerodynamics, electromagnetism …)
- In a wide variety of Applications in industry, in particular in numerical **simulation**

Electricity (power grids)



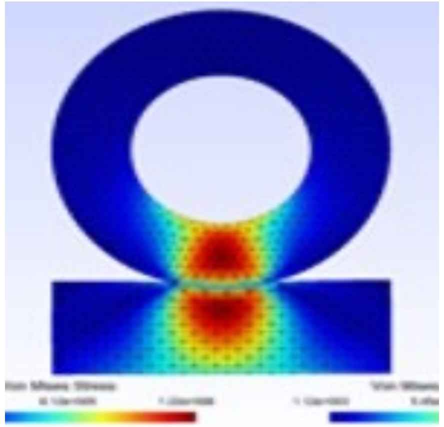
Picture from Marot, A., et al. (2018). Guided machine learning for power grid segmentation. In *2018 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe)* (pp. 1-6).

Aerodynamics



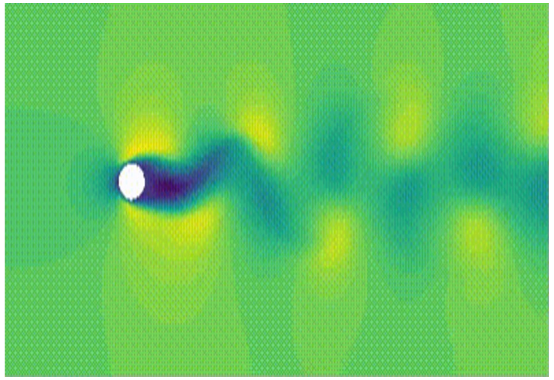
Merino-Martínez *et al. CEAS Aeronautical* Journal (2019).

Solid Mechanics pneumatics



From the internet

Fluid Flows/Dynamics



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

Domain Challenges : Physical systems that are
- Complex to model/solve analytically
- Compuationally expensive to solve numerically

eg. , Computational Fluid Dynamics – CFD, Turbulance, Flows

Scientific Challenges

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

Hybrid modeling: combining Machine Learning and ***Physics***

Hybrid modeling: combining ML and *Physics*

- ➔ Enables prior scientific knowledge based on physics to be taken into account in data-driven machine learning methods: e.g approcahes include 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 …

In engineering, it allows

- ➔ the integration of analytical knowledge from physical laws governing the studied engineering systems

- to augment th statistical knowledge learned from observed/measured data (eg. Information extracted by deep learning from data)

- for reducing the high cost of physical simulation, in particular in the industrial sector

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

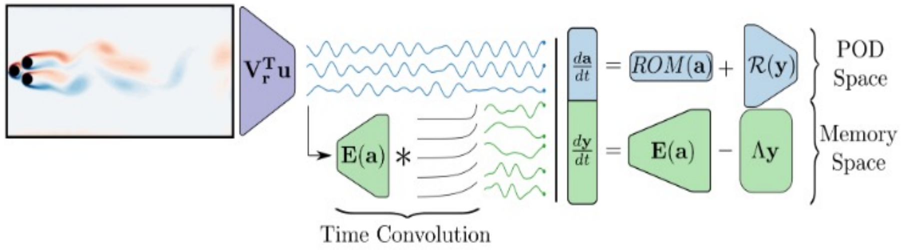
HSA Project : Simulation/machine learning hybrid modeling

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

➔ Hybrid Machine Learninrg as surrogate models for physical simulation, aiming to

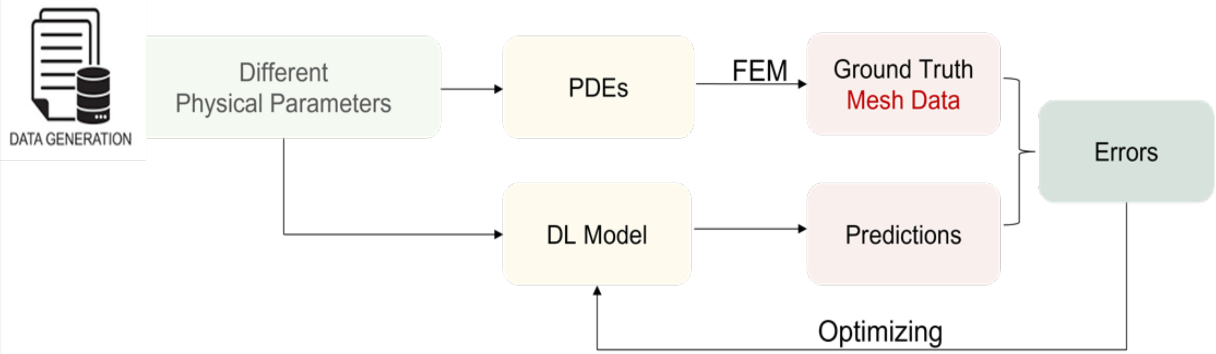
Replace physical solvers with



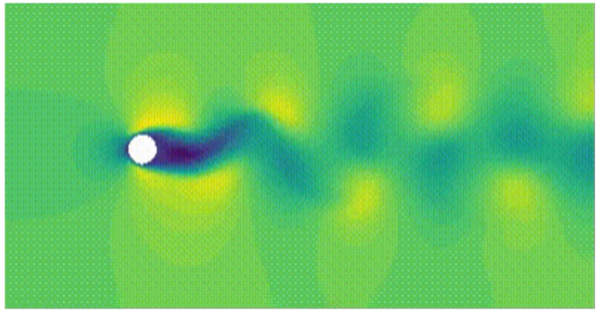
High-Dimensional non-linear Physical Equations

➔ Deep learning intergrating physical constraints (eg. Deep Graph Nets for PDEs)

➔ Deal with high-dimensional, non-linear, and complex structurs (e.g reduced modeling, ..)



Deep Graph Neural Networks for Numerical Simulation of PDEs. PhD of W. Liu. 2023 (LISN, Inria/SystemX). Read Online



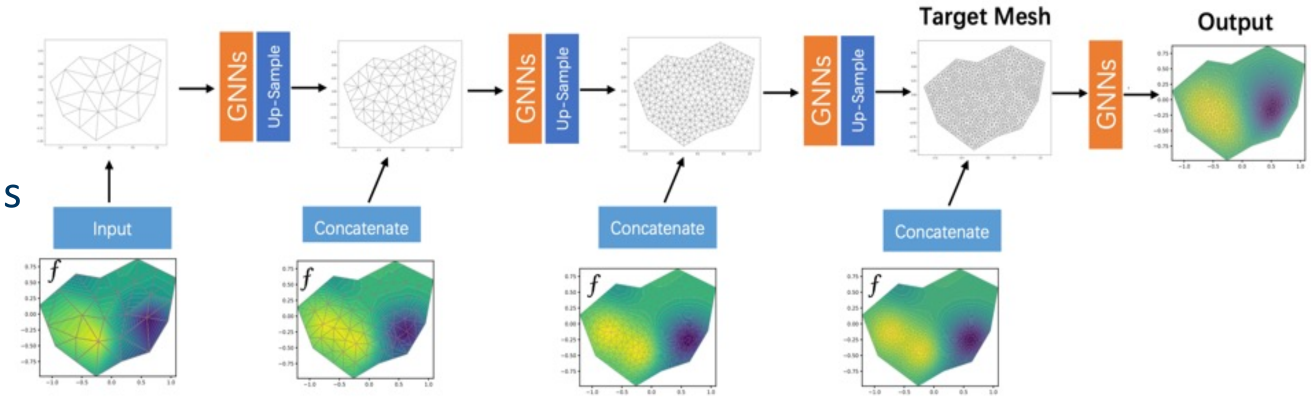
Reduced models and deep learning for PDEs
PhD Thesis of E. Menier, 2024 (LISN, Inria/SystemX)

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

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

Project HSA : simulation and deep learning of graphs

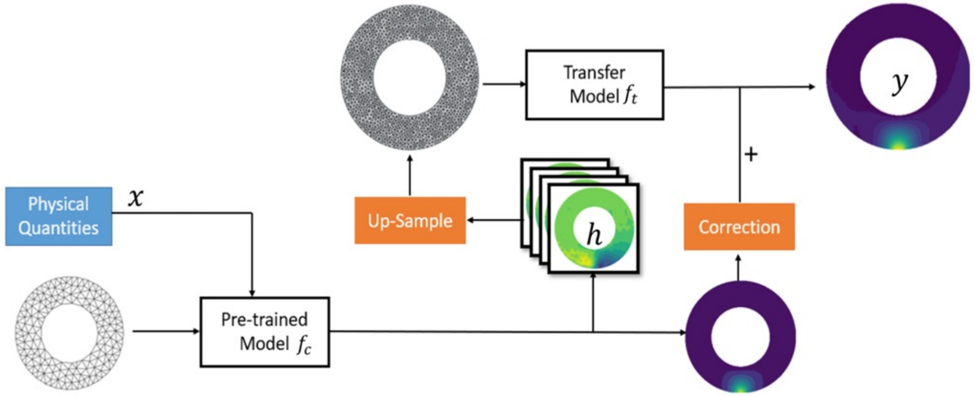
Graph Neural Nets for 3D **meshes**



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



Wheel contact profile

Physics: contact equations

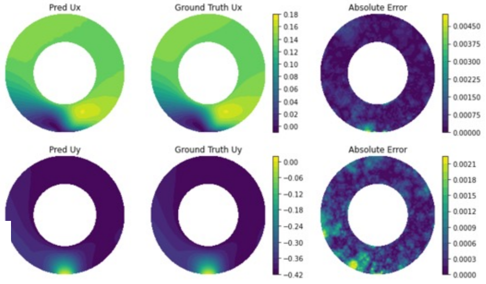
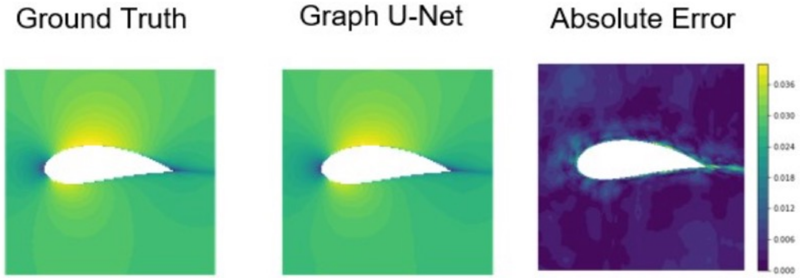


Figure 6.3: An example of wheel contact prediction

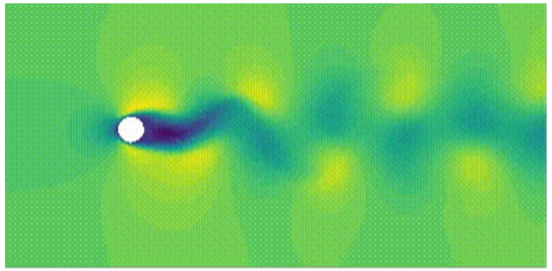
Prediction of the **airflow** profile around an aircraft wing (Air Foil)

Physics: Navier-Stokes equations



PhD theis of W. Liu, 2023 (LISN, Inria/SystemX)

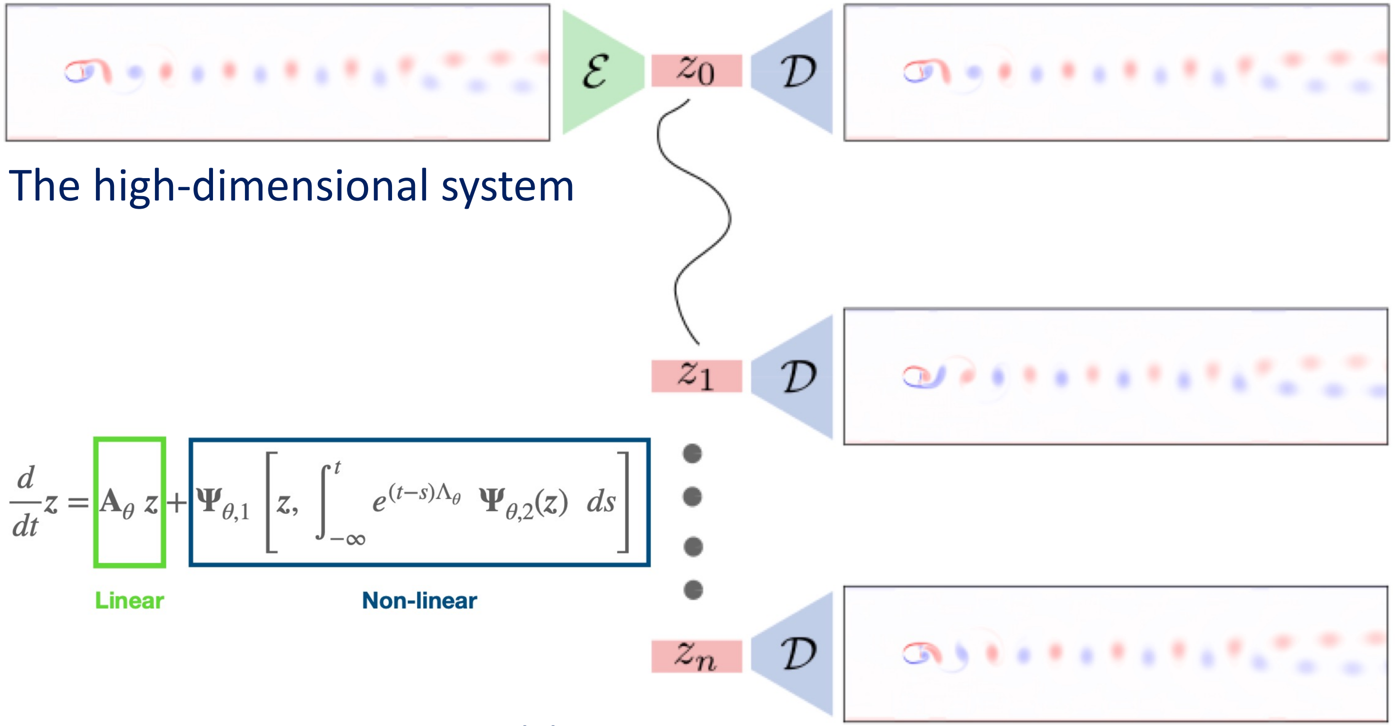
Dynamics: Hybrid ML for HD dynamical physical systems



High-Dimensional non-linear Dymical Systems:

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

Interpretable learning of effective dynamics (ILED) architecture:



The high-dimensional system

$$\frac{d}{dt}z = \underbrace{\boxed{\mathbf{A}_\theta\, z}}_{\text{Linear}} + \underbrace{\boxed{\mathbf{\Psi}_{\theta,1}\left[z,\ \int_{-\infty}^{t} e^{(t-s)\Lambda_\theta}\ \mathbf{\Psi}_{\theta,2}(z)\ ds\right]}}_{\text{Non-linear}}$$

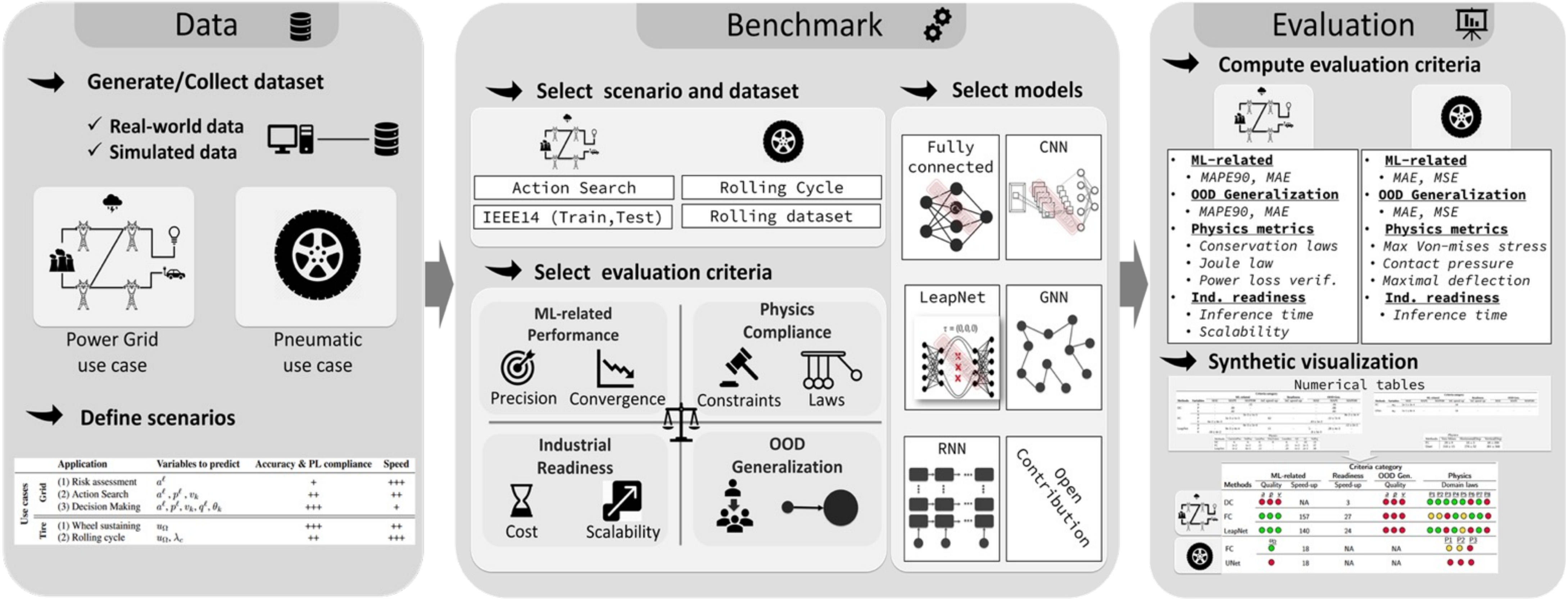
The lower-dimensional representation (**z**) is propagated in time using a linear and a non-linear part based on the Mori-Zwanzig formalism

The decoder D reconstructs the high-dimensional systems.

Menier, E., et al. (2023). Interpretable learning of effective dynamics for multiscale systems. *arXiv preprint arXiv:2309.05812*. Read Online
PhD Thesis of E. Menier, 2024 (LISN, Inria/SystemX), 2024

LIPS: Platform of validation of hybrid AI models

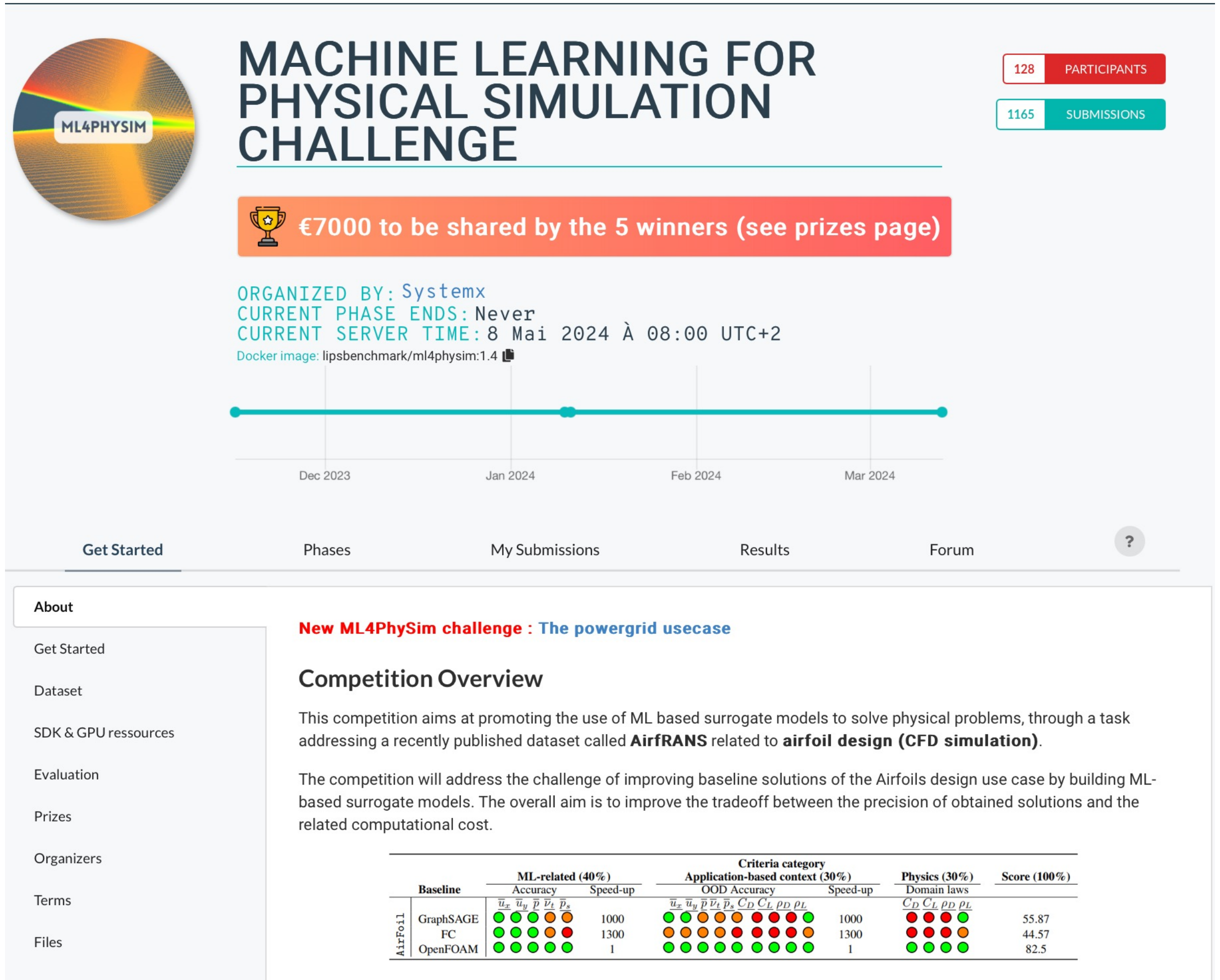
- **LIPS** : Learning Industrial Physical Simulation benchmark suite *(Result of the project HSA-IA2)*
- Evaluation of physical simulator augmented by machine learning
- Open-source Framework https://github.com/IRT-SystemX/LIPS Published at NeurIPS2022
- 1st framework for evaluating augmented physical simulators
- 7 use cases integrated



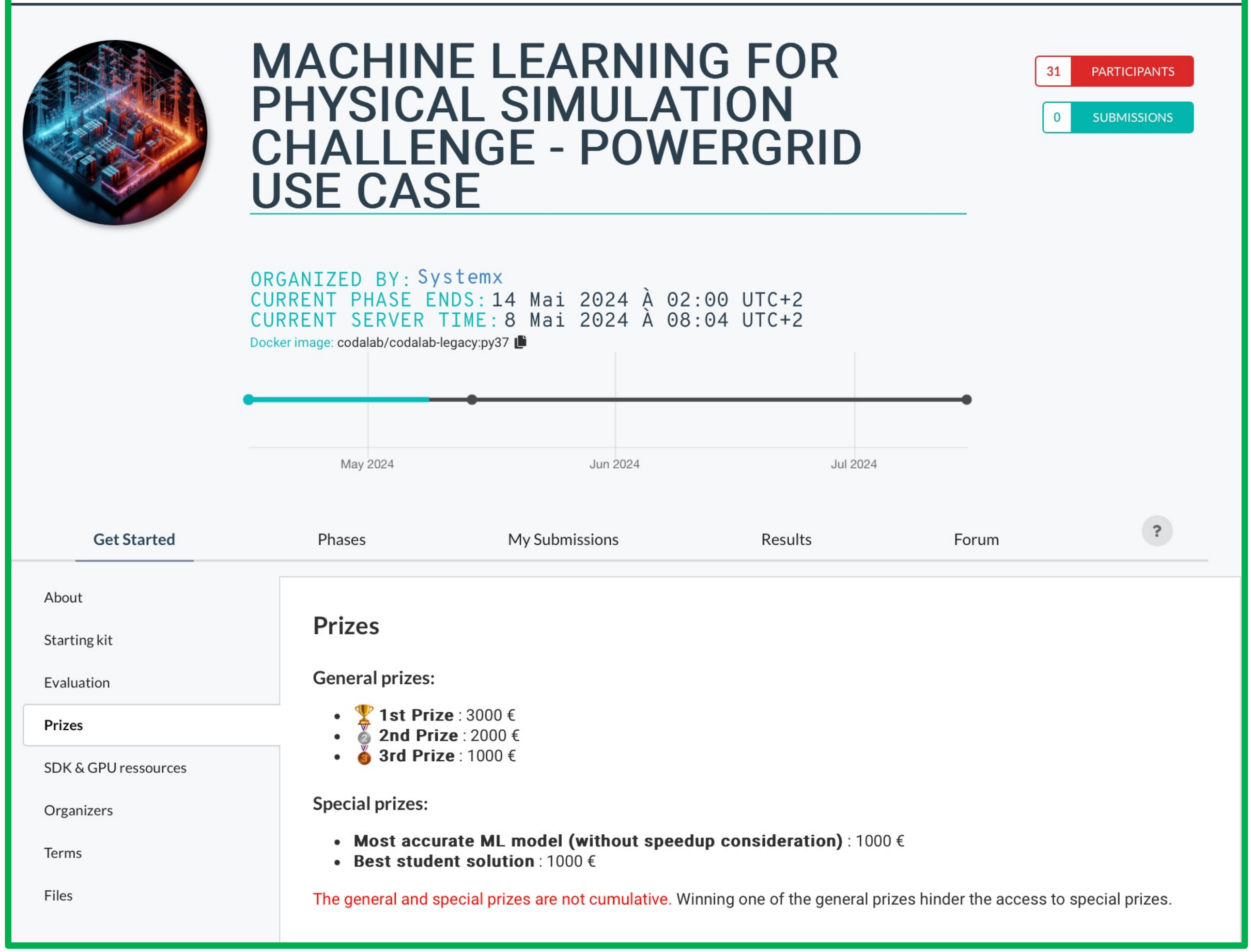
Competitions on Codabench/codalab

LIPS hosts the two following competitions:

https://www.codabench.org/competitions/1534/ (Closed)



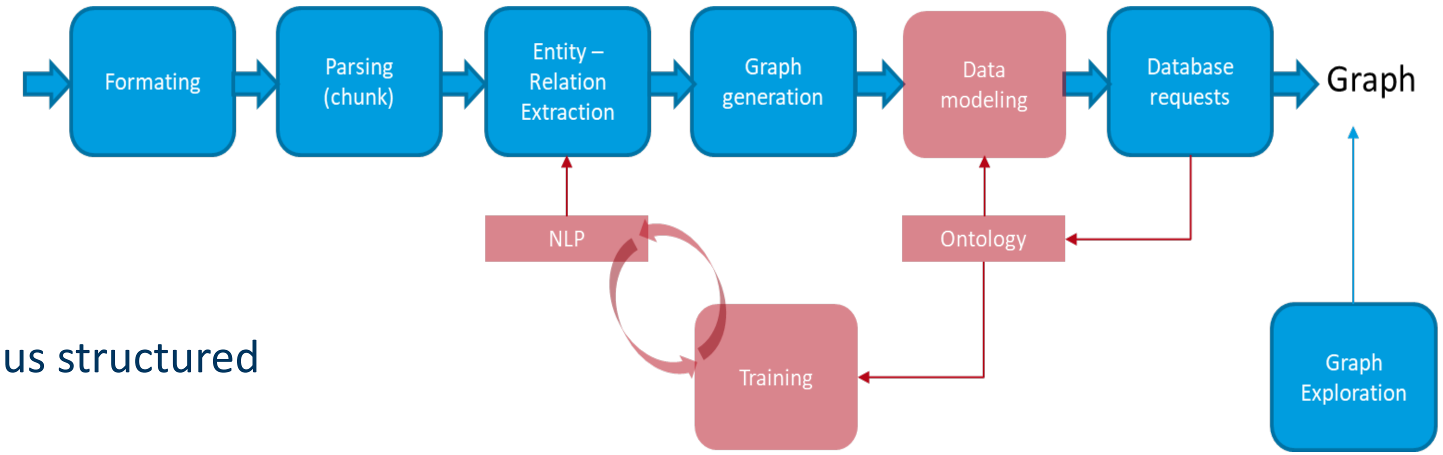
https://www.codabench.org/competitions/2378/ **Running!**



Project SMD : Integrating Expert/Business *semantics* in ML

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

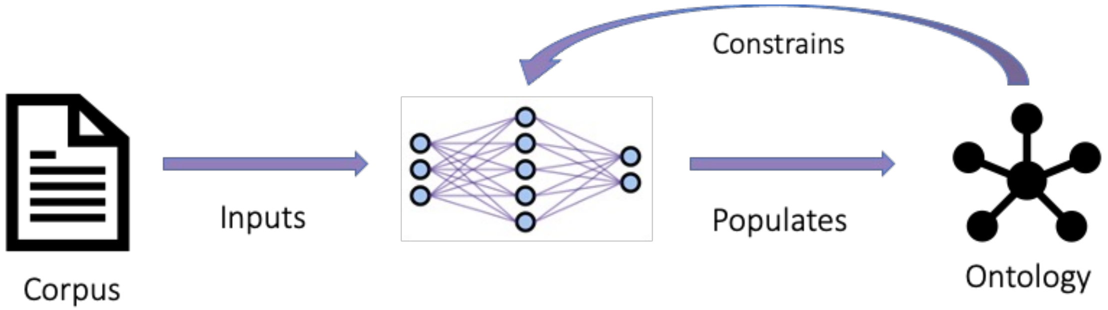
Framework:
Building ML-based on knowledge graphs from expert/business language data



Approach

- 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

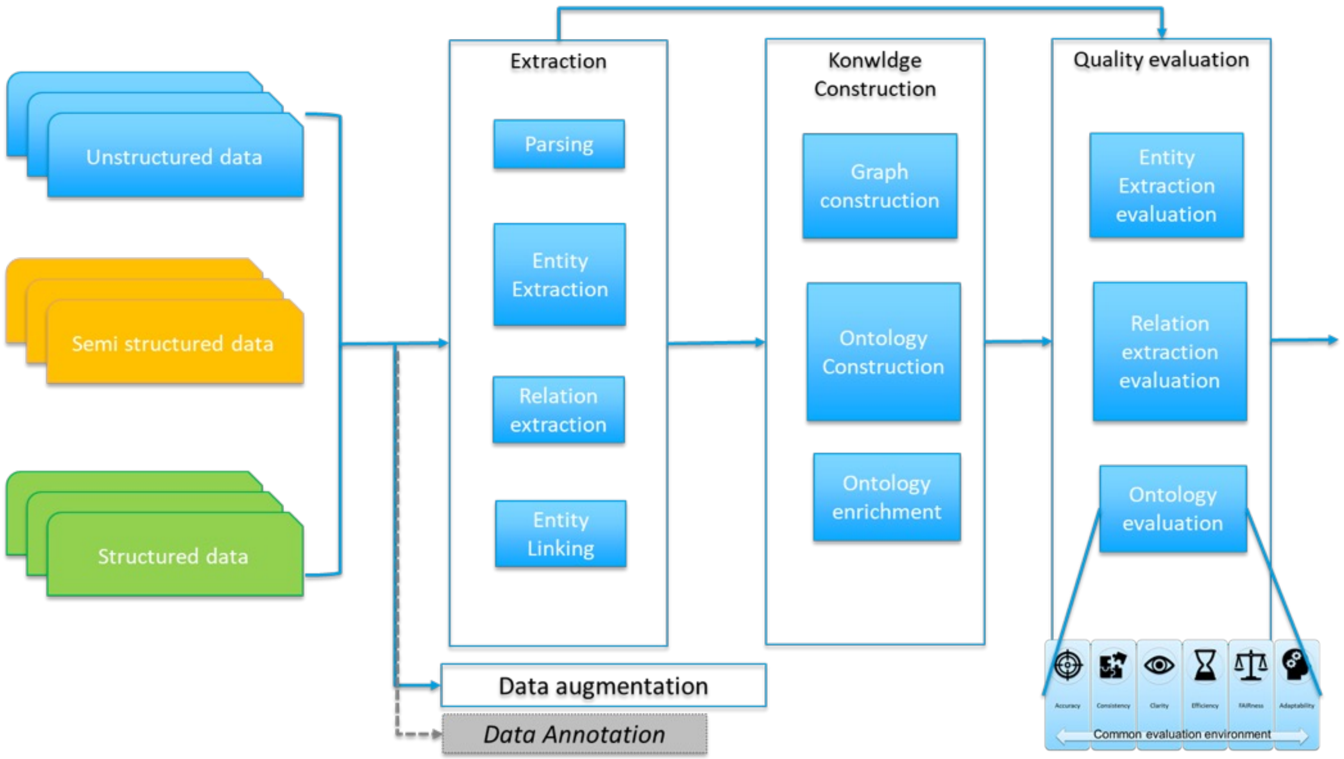
Neuro-symbolic pipeline

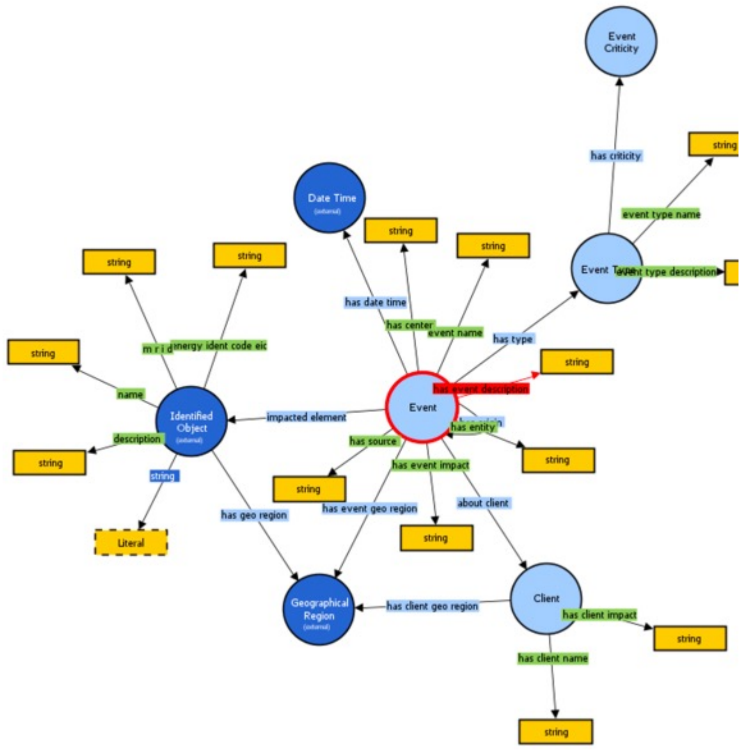


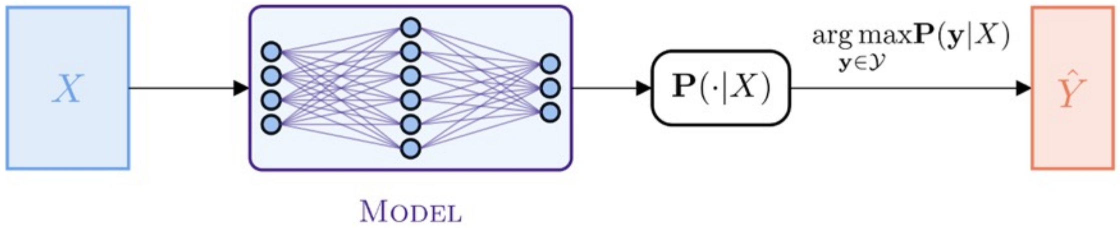
IA² Intelligence artificielle et ingénierie augmentée

SMD Project

Project SMD : Integrating Expert/Business semantics in ML







=> Insertion of logical rules at the inference step :

- Recalculer P(y|X, α) in lieu of the learnt P(y|X)
- α is a rule encoding the validity of the prediction $\hat{y}$

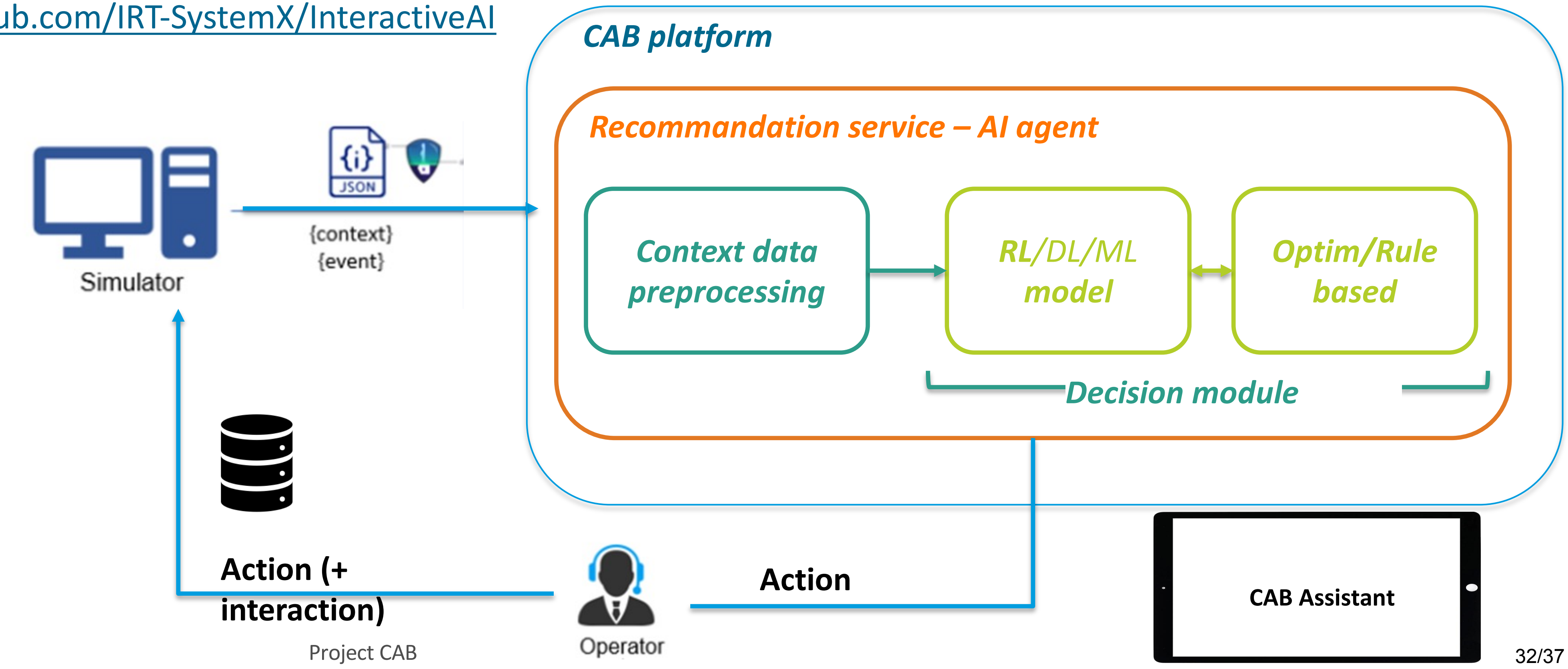
Ontologies and (machine/transfer) learning for multimedia document analysis. PhD thesis of A. Ledaguenel (in progress, MICS/SystemX)

Human – AI Interaction: CAB Project (Bi-directional assistant)

Objectives and challenges

- Design and implement a Bidirectional Assistant to support operators in network supervision and aircraft piloting activities
- Bidirectional assistant: The assistant can learn from and to (inform) the operator
- **Platform**: https://github.com/IRT-SystemX/InteractiveAI

Project IA² CAB



Outline

Challenges and industrial applications in

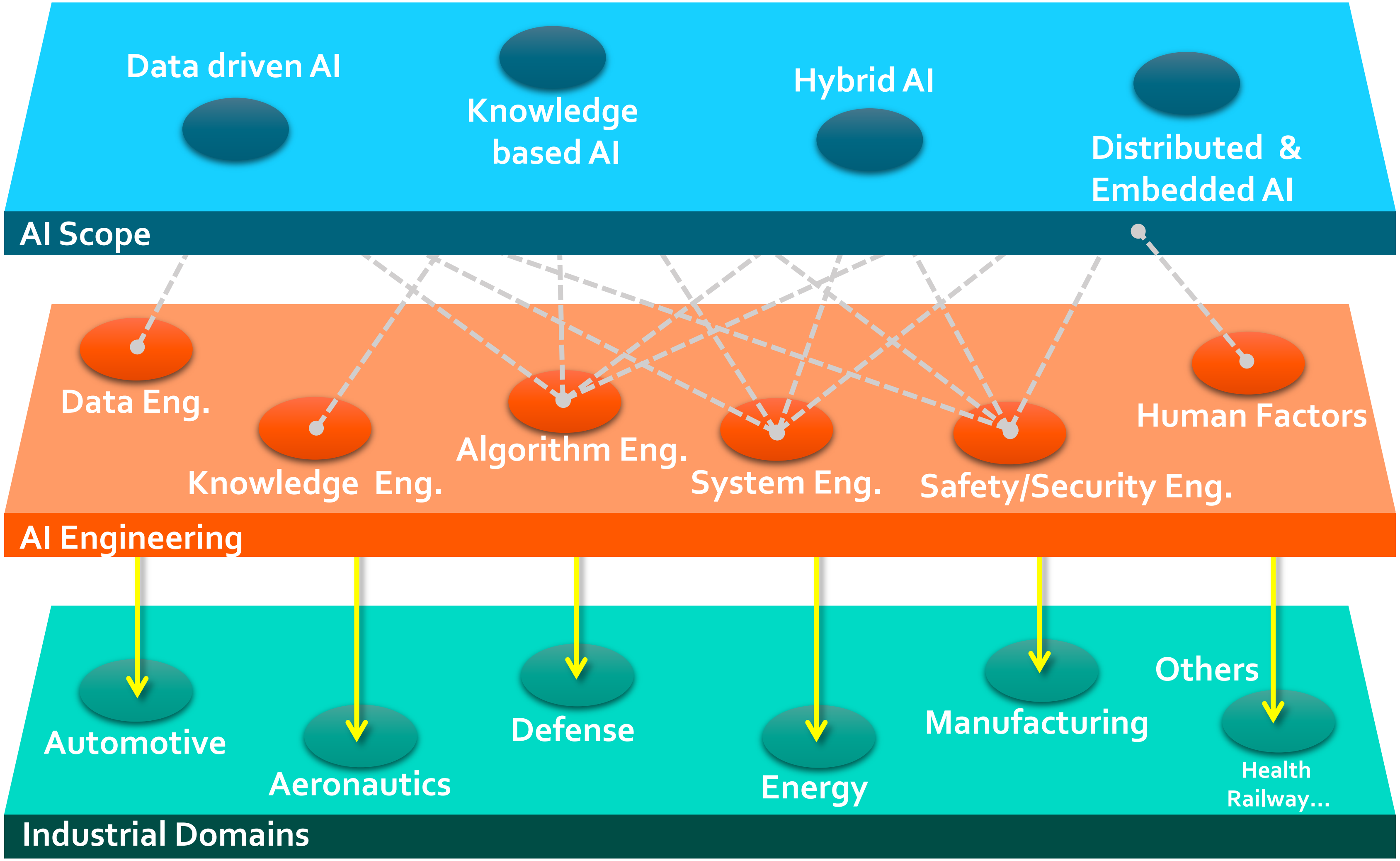
- Hybrid AI: exploiter les connaissances métiers de natures physique et symbolique dans les modèles d’apprentissage

- Trustworthy AI: concevoir et industrialiser des systèmes à base d’intelligence artificielle de confiance

Challenges and industrial applications in Trustworthy AI: Confiance.AI programme

A French unique community to design and industrialise trustworthy AI-based critical systems

Multi-technology, multi-domain, multi-engineering



Confiance.ai : Key figures

4 YEARS	Duration
9	Large industrial groups
4	Research centers
7	Thematic projects
+300	FTP involved over 4 years
2	Sites: Paris-Saclay and Toulouse
45 M€	Budget
12	Associate partners (laboratories, SMIs, startups)

Scientific challenges as part of the ongoing PhD theses within confiance.ai

- **confiance.ai:** methods and tools for trusted AI ➔ High expectations for industry
- In parallel of development of tool chain, many scientific challenges remain, including:
- Ongoing PhD theses within confiance.ai :
  - PhD Theis of Adrien Le Coz: Data coverage and operational **domain design** ODD (Computer Vision)
  - PhD Thesis of Paul La Barbarie: **Robustness** to ‘patch’ adversarial attacks (Computer Vision)
  - PhD Thesis of Lucas Schott: **Reinforcement** learning and human in the loop
  - PhD Thesis of Gayane Taturyan: Statistical control of **Fairness**/Bias in Machine Learning (Stat)
  - PhD Theis of Houssem Ouertatani : **Optimized** hardware **deployment** (Neural Architecture Search)
  - PhD Thesis of Abdelmouaiz Tebjou: Conformal prediction (deployed) AI algorithms **monitoring**

SystemX
INSTITUT DE RECHERCHE TECHNOLOGIQUE

Accélérateur de la transformation numérique

FRANCE 2030

THANK YOU FOR YOUR ATTENTION