

DANI 23 mai 2023

■ Chaire Industrielle ANR DANI

Données Apprentissage et modèles Numériques pour l'Ingénieur

David Ryckelynck

MINES Paris - PSL Université, France

Durée : 4 ans

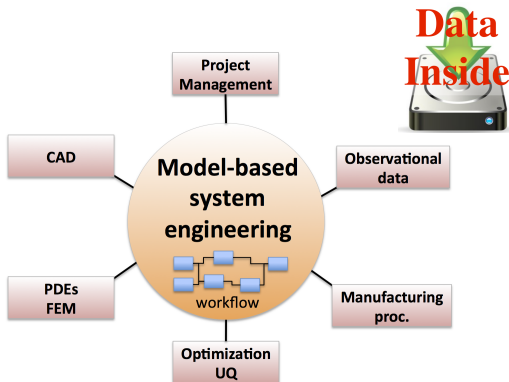
Démarrage envisagé octobre 2024, dépôt en mars 2024

Objectifs : 5 thèses de doctorat, 4 stages de mastère, 4 postdocs, 4 industriels partenaires

Domaine de la chaire : AI for science and engineering

Industrial fields :

- Nuclear industry
- Aeronautics industry
- Automotive industry
- Defence industry
- Energy industry
- Manufacturing industry
- ...



**Ingénierie augmentée
par l'IA**

Avec l'IA, trouver de nouvelles marges de conception, de production, d'exploitation, de réduction de déchets...

Sans rien oublier en science et technologie.

IA Hybride pour une ingénierie augmentée

Définition : Du machine learning à l'état de l'art, avec :

- Des données structurées : images, tenseurs, graphes, ou modèles physiques
- ou, apprentissage automatique avec données synthétiques issues de la modélisation à base physique,
- ou, des Loss contraintes par les connaissances scientifiques et techniques (par exemple PINN),
- ou, des méthodes numériques à base physique pour la validation de prévisions par réseaux de neurones,
- ou, du manifold learning pour la modélisation (réduction de modèle, sous modèles, CL, CI...).
- ou, augmentation de données par jumeaux numériques à haute fidélité,
- ou, Data Pruning en conservant des capacités de modélisation.

Exemples d'applications ciblées

- Maintenance prédictive, dérive de process/procédés
- Vision par ordinateur pour l'industrie, contrôle non destructif en ligne, détection d'anomalies (sans exemples suffisants d'anomalies),
- Jumeaux numériques issus de données structurées, avec ou sans variabilité géométrique,
- Apprentissage de variétés (espaces latent), apprentissage auto-supervisé, réduction de dimension,
- Propagation d'incertitudes accélérée, passer des observations aux prévisions pour la conception ou l'exploitation de systèmes,
- Données et modèles multimodaux pour les matériaux dans les systèmes,
- Augmentation/élagage de données technologiques et scientifiques,
- Apprentissage fédéré sur procédés de fabrication/diagnostic communs mais avec des données privées (ex. soudage, fab. add., assemblage...)
- ...

Résultats attendus

Objectifs de résultats :

- Une communauté d'ingénieurs, d'étudiants, de chercheurs.
- Des méthodes et des experts en Ingénierie Augmentée par l'IA.
- Des brevets, des thèses de doctorats, des articles de conférence et de journaux internationaux.
- Des "briques" python, pour appliquer les méthodes d'apprentissage à de l'ingénierie augmentée par de l'IA (utilisation au maximum de l'existant, Mordicus, sklearn, PyTorch, ect.)

Types de projets de recherche de la chaire DANI :

- P1 : Stage de mastère ou de master de 6 mois avec pré-publication d'un article
- P2 : Mission post-doctorale de 12 à 18 mois avec publication d'un à deux articles
- P3 : Thèse de doctorat en 3 ans avec publication d'au moins deux articles

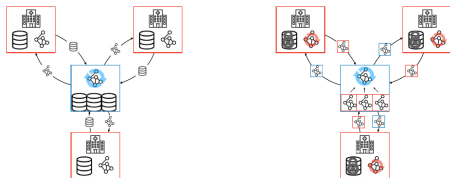
Exemple de sujet de thèse potentiel :

Federated machine learning for anomaly detection in additive manufacturing

D. Ryckelynck

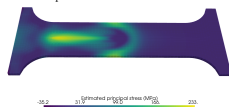
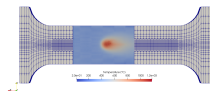
Travaux antérieurs à Mines Paris
PSL :

- Mission de l'option IDSC chez Google en 2020
- Stage d'Arthur Pignet en 2022 chez Owkin



a / Classic ML on pulled data.

b / Cross-silo FL.



-35.2 31 233
Estimated principal stress (MPa)

Pablo P. Alvarez's PhD 2022 (EDF)

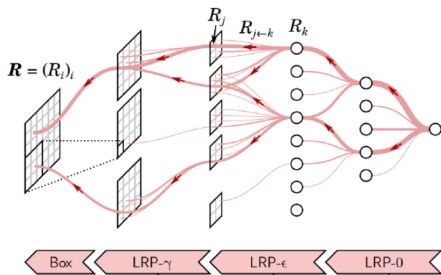
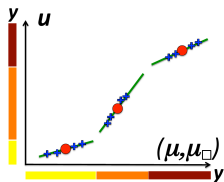
Un **consortium de détenteurs de données** entraîne conjointement un modèle sur des **données sensibles** tout en **minimisant l'exposition des données**.

Contrairement aux scénarios classique où les données de tous sont centralisées. en FL les données restent dans le serveur de chaque propriétaire des données.

Exemple de sujet de thèse potentiel :

Manifold Learning for fast and explainable digital twins under geometrical uncertainties

D. Ryckelynck



Layer-wise Reliance Propagation [Bach 2015 : Propagate the output of the network backwards by means of propagation rules.

- Ambient space for geometrical variations
- Manifold learning of latent spaces (sparse AE + deep classifier **GNN**)
- Explainable classifier
- Multimodal data augmentation

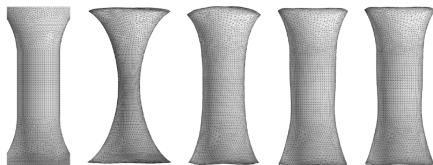


FIGURE 4.5 – Illustration of the CAD mesh projection along several modes (from left to right) : all modes, modes 2, 10, 100, 1000.

Axel Aublet's PhD thesis 2023 (SafranTech, Mines Paris PSL)

Exemple de sujet de thèse potentiel :

Graph Matching for digital twining via ROM-nets [Thomas Daniel 2020]

Le Graph Matching a un énorme potentiel de développement pour traiter les matériaux et les fluides dans les systèmes.

En effet, tout modèle numérique est assimilable à un graphe. Les conditions limites, l'état initial, les dommages, les interfaces..., peuvent être considérés comme des labels sur des graphes.

On peut associer à cela des méthodes de federated learning, pour mettre en oeuvre le manifold learning.

Autres exemples de sujets de thèse potentiels

■ Données

- Federated learning for image-based anomaly detection on technical or rough surfaces
- Manifold learning of as-manufactured shapes for uncertainty quantification in design steps using PDEs / functional metrology, ROM-net for graphs
- Optimal search for missing data or for missing knowledge

■ Apprentissage hybride

- Learning real-time crash simulations using Hamiltonian equations
- Learning time series predictions on spatially structured data
- Model order reduction for contact mechanics in multibody and deformable systems
- Predictive maintenance augmented by machine learning for networks of elongated structures (rails, pipeline, beams,...), tensor products in convolutions
- Scientific verification of AI-outputs

■ Architectures

- Graph Neural Networks for engineering applications under conservation laws
- AI-accelerated multiscale material models : towards streamlined data acquisition
- Deep learning of tensor decompositions for high dimensional convolutions, application to multiphysics learning
- Multiscale learning on structured data for multiphysics engineering in complex systems
- Multimodal image-based digital twings involving time series for multiphysics systems
- Geometric learning for digital twining in industry 4.0

Rejoignons nous sur des cas d'usage industriels

Description de chaque cas d'usage :

- Confidentialité : oui/non (NDA)
- Titre non confidentiel
- Une à deux images de l'application
- Contexte 100 mots maximum
- Type de données (observées, synthétiques, publique, confidentielles, images, tenseurs, graphe,...)
- Nombre d'instances disponibles (> 30) et dimension des données (> 100)
- Objectif industriel visé en 200 mots maximum
- Proposition de tâches de ML à développer : régression, classification, extraction de caractéristiques, visualisation, accélération du traitement, réduction de modèle, reinforcement learning, Federated Learning, XAI ...

Extraits de liste de publications

- Deep multimodal autoencoder for crack criticality assessment, Hugo Launay , David Ryckelynck , Laurent Lacourt , Jacques Besson , Arnaud Mondon et al., International Journal for Numerical Methods in Engineering, 2022
- Uncertainty quantification for industrial numerical simulation using dictionaries of reduced order models, Thomas Daniel , Fabien Casenave , Nissrine Akkari , David Ryckelynck , Christian Rey et al., Mechanics & Industry, 2022
- Multimodal data augmentation for digital twining assisted by artificial intelligence in mechanics of materials, Axel Aublet , Franck N'Guyen , Henry Proudhon , David Ryckelynck, Frontiers in Materials, 2022
- Deep learning model to assist multiphysics conjugate problems, George El Haber , Jonathan Viquerat , Aurelien Larcher , David Ryckelynck , Jose Alves et al., Physics of Fluids, 2022
- An updated Gappy-POD to capture non-parameterized geometrical variation in fluid dynamics problems, Nissrine Akkari , Fabien Casenave , David Ryckelynck , Christian Rey, Advanced Modeling and Simulation in Engineering Sciences, 2022
- A Bayesian Nonlinear Reduced Order Modeling Using Variational AutoEncoders, Nissrine Akkari , Fabien Casenave , Elie Hachem , David Ryckelynck, Fluids, 2022
- Condition number and clustering-based efficiency improvement of reduced-order solvers for contact problems using Lagrange multipliers, Simon Le Berre , Isabelle Ramiere , Jules Fauque , David Ryckelynck, Mathematics, 2022
- Uncertainty quantification in a mechanical submodel driven by a Wasserstein-GAN, Hamza BOUKRAICHI , Nissrine Akkari , Fabien Casenave , David Ryckelynck, IFAC-PapersOnLine, 2022
- Real-Time Data Assimilation in Welding Operations Using Thermal Imaging and Accelerated High-Fidelity Digital Twinning, Pablo Pereira Álvarez , Pierre Kerfriden , David Ryckelynck , Vincent Robin, Mathematics, 2021
- A modular U-Net for automated segmentation of X-ray tomography images in composite materials, Joao P C Bertoldo , Etienne Decenci re , David Ryckelynck , Henry Proudhon, Frontiers in Materials, 2021
- Model order reduction assisted by deep neural networks (ROM-net), Thomas Daniel , Fabien Casenave , Nissrine Akkari , David Ryckelynck, Advanced Modeling and Simulation in Engineering Sciences, 2020

Poursuivons le débat ...

... et préparons le détail des axes thématiques de la chaire.