



IMT Mines Alès
École Mines-Télécom

LA SCIENCE & LA CRÉATIVITÉ POUR INVENTER UN MONDE DURABLE

DOCTORAT EN APPRENTISSAGE AUTOMATIQUE

Etablissement	IMT Mines Alès (Ecole Nationale Supérieure des Mines d'Alès)
Affectation principale	Centre d'Enseignement et de Recherche en Informatique et Systèmes (CERIS)
Résidence administrative	Alès (Département du Gard – Région Occitanie)
Date de prise de poste	01/10/2024

1. IMT et IMT Mines Alès

L'institut Mines-Télécom (IMT), grand établissement au sens du code de l'éducation, est un établissement public scientifique, culturel et professionnel (EPSCP) placé sous la tutelle principale des ministres chargés de l'industrie et du numérique. Premier groupe d'écoles d'ingénieurs en France, il fédère 11 écoles d'ingénieur publiques réparties sur le territoire national, qui forment 13 500 ingénieurs et docteurs. L'IMT emploie 4500 personnes et dispose d'un budget annuel de 400M€ dont 40% de ressources propres. L'IMT comporte 2 instituts Carnot, 35 chaires industrielles, produit annuellement 2100 publications de rang A, 60 brevets et réalise 110M€ de recherche contractuelle.

Créé en 1843, IMT Mines Alès compte à ce jour 1400 élèves (dont 250 étrangers) et 380 personnels. L'école dispose de 3 centres de recherche et d'enseignement de haut niveau scientifique et technologique, qui œuvrent dans les domaines des matériaux et du génie civil (C2MA), de l'environnement et des risques (CREER), de l'intelligence artificielle et du génie industriel et numérique (CERIS). Elle dispose de 12 plateformes technologiques et compte 1600 entreprises partenaires.

2. Projet de recherche

Titre: "Apprentissage de représentations neuro-symboliques désenchevêtrées à partir de graphes de connaissances multimodaux"

Mots-clef: Apprentissage profond, Apprentissage de représentations interprétables, Multimodalité, Graphes de connaissances, Apprentissage neuro-symbolique

Résumé: L'objectif du projet est la conception d'une méthodologie d'apprentissage de représentations multimodales interprétables au travers d'un désenchevêtrement des données par un graphe de connaissances (disentangled representations) liant les informations conceptuelles à des données de différentes modalités associées au concept (mentions du concept dans du texte, occurrences dans des images, des vidéos, des séquences de poses de mouvement humain et de l'activité cérébrale ou physiologique). Plutôt que d'approcher le problème au travers d'une méthodologie bout-en-bout partant d'un grand jeu de données et entraînant un modèle spécifique, nous souhaitons partir de modèles dits « fondation » existants pour les modalités (Large Language Models, Pretrained Vision Transformers, etc.) et du graphe de connaissances au travers de techniques de plongement de graphes, et calculer une représentation contrastive commune permettant la confrontation des différentes modalités. Cette approche vise à exploiter au mieux ces grands modèles de référence afin de minimiser le coût et l'impact du calcul des représentations multimodales, mais présente aussi l'avantage de

permettre un calcul à la volée sur des données liées ouvertes (linked open data) par exemple sur la base de sous-graphes émanant de requêtes SPARQL. Ce projet combine deux volets de recherche distincts portant sur le désenchevêtrement des représentations par les connaissances d'une part, et l'apprentissage multimodal contrastif interprétable, tout en proposant d'étendre ce dernier paradigme à plus de deux modalités ($n > 2$). Cas applicatifs possibles en Industrie 5.0 ou en IA pour la santé en lien avec le mouvement humain et de la neuroimagerie.

Références bibliographiques :

M. Lee and V. Pavlovic, "Private-Shared Disentangled Multimodal VAE for Learning of Latent Representations," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, Nashville, TN, USA, 2021, pp. 1692-1700, doi: 10.1109/CVPRW53098.2021.00185.

S. Zhang, X. Rao, Y. Tay, and C. Zhang. 2021. Knowledge Router: Learning Disentangled Representations for Knowledge Graphs. In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 1–10, Online. Association for Computational Linguistics.

X. Wang, H. Chen, and W. Zhu. 2023. Disentangled Representation Learning for Multimedia. In Proceedings of the 31st ACM International Conference on Multimedia (MM '23). Association for Computing Machinery, New York, NY, USA, 9702–9704. <https://doi.org/10.1145/3581783.3613859>

J. Chen and A. Zhang. 2024. On Disentanglement of Asymmetrical Knowledge Transfer for Modality-Task Agnostic Federated Learning. Proceedings of the AAAI Conference on Artificial Intelligence, 38(10), 11311-11319. <https://doi.org/10.1609/aaai.v38i10.29010>

3. [Encadrement](#)

Centre de Recherche et d'enseignement : Centre d'Enseignement et de Recherche en Informatique et Systèmes (CERIS)

Ecole doctorale : ED n°166 - I2S (Information, Structures et Systèmes)

Equipe d'encadrement : Andon Tchechmedjiev, Sebastien Harispe

4. [Profil recherché](#)

Le candidat idéal a un profil data science / deep learning avec une coloration ingénierie de connaissances et/ou neurosciences et/ou linguistique et/ou vision par ordinateur, ou toute permutation de ces domaines. Ce dernier devra posséder une ouverture d'esprit vers d'autres disciplines et montrer une capacité à intégrer de nouveaux concepts relevant d'autres paradigmes épistémologiques, mais également un sens de l'empathie et du travail d'équipe développé.

5. [Contacts](#)

- ▶ Sur le projet de recherche : andon.tchechmedjiev@mines-ales.fr
- ▶ Sur les aspects administratifs : Anne-Catherine Denni (anne-catherine.denni@mines-ales.fr)