

# Omar AbdelKader

Lille, France



---

## À propos de moi

Je suis chercheur en IA, avec une solide expérience en intelligence artificielle et en génie logiciel.  
Je prépare actuellement un doctorat en informatique à l'Université de Lille.

---

## Responsabilités

- Fondateur et mainteneur de l'organisation **Pharo-LLM**.
- Membre du comité d'organisation, GDR-GPL'26.
- Membre du comité d'organisation, **ESUG'26**, **ESUG'25**, **ESUG'24**.
- Sous-relecteur pour **BENEVOL'25**.
- Chef de projet et mainteneur de plusieurs projets (**ChatPharo**, **Pharo-Copilot**, **Pharo-Infer...**).
- Mentor dans **GSoC'26**, **GSoC'25**.
- Président de **NeuroTech-Lille**, une association étudiante axée sur l'IA et les neurosciences.
- Contributeur au **Ministère de la Transition Écologique** (*Club IA et Transition écologique*).
- Membre de l'organisation **Pharo-AI**.
- Contributeur au **Conseil National du Numérique (CNNum)** (*Café IA*).
- Encadrement de nombreux projets logiciels et IA aux niveaux Licence et Master, incluant des projets orientés recherche ainsi que des projets appliqués en industrie.

---

## Publications

- AbdelKader, O., Ducasse, S., Zaitsev, O., Robbes, R., & Polito, G. (2025). Package-Aware Approach for Repository-Level Code Completion in Pharo. *International Workshop on Smalltalk Technologies (IWST)*. DOI.

---

## Expérience

**Chercheur en IA — INRIA (Villeneuve-d'Ascq, France) oct. 2024 – oct. 2027**

- Amélioration de l'autocomplétion et de la génération de code à l'aide de LLMs, en ciblant spécifiquement le langage de programmation Pharo, qui dispose de données d'entraînement limitées.
- Développement de techniques d'autocomplétion, d'inférence de types et de déploiement dans l'IDE de Pharo, avec un focus sur les performances à l'exécution.

**Ingénieur IA — INERIS (Verneuil-en-Halatte, France) sept. 2023 – sept. 2024**

- Développement de « INERIS-IA », un outil de classification de documents textuels selon les objectifs stratégiques de l'INERIS, en mobilisant des techniques de ML et de NLP.
- Conception de requêtes booléennes pour la recherche documentaire et amélioration de la qualité du corpus via la similarité de documents et l'extraction de mots-clés.

**Stage — Chercheur en IA — LIPN (Villetaneuse, France) juin 2023 – août 2023**

- Comparaison de techniques d'échantillonnage pour la planification probabiliste, notamment pour générer des narrations littéraires.
- Évaluation de méthodes incluant le Score Function Estimator (SFE) et des techniques avancées telles que Gumbel-Softmax, en mesurant leur efficacité à produire des histoires cohérentes et créatives.

**Développeur logiciel — BEON-IT (Beyrouth, Liban) mai 2022 – juin 2022**

- Java · Design Patterns · .NET Framework · Threading · Microsoft SQL Server

## Développeur Full-Stack — SSCC-IT (Andket, Liban) *nov. 2020 – avr. 2021*

- PHP & SQL · Design Patterns · Microsoft SQL Server

---

## Formation

### Université de Lille — Villeneuve-d'Ascq, France

*Doctorat en informatique* (oct. 2024 – Présent)

### Université de Lorraine — Nancy, France

*Master en Traitement Automatique des Langues (NLP)* (sept. 2022 – sept. 2024)

### Université libanaise — Beyrouth, Liban

*Licence en Data Science* (sept. 2019 – juil. 2022)

---

## Logiciels

### Pharo-Copilot

- Outil intelligent d'autocomplétion et de génération de code pour Pharo, inspiré de GitHub Copilot.
- Axé sur l'autocomplétion à l'échelle des dépôts et des packages.
- Exploitation de LLMs pour améliorer la productivité des développeurs dans un langage à typage dynamique.
- **GitHub Repository:** Pharo-Copilot

### ChatPharo

- Interface conversationnelle en direct entre les développeurs Pharo et les grands modèles de langage.
- Permet des échanges interactifs avec des LLMs directement depuis l'environnement Pharo.
- Conception d'une architecture ouverte et extensible pour expérimenter plusieurs backends de LLMs.
- **GitHub Repository:** ChatPharo

### INERIS-IA

- Plateforme IA développée à l'INERIS pour la classification de documents et la gestion des connaissances.
- Interface web basée sur Flask intégrant plusieurs modèles de ML et de NLP.
- Entraînement de modèles de classification alignés sur les objectifs COP 2027 et les thématiques stratégiques de l'INERIS.
- Mise en place de fonctionnalités supplémentaires : similarité de documents, extraction de mots-clés et amélioration de la qualité du corpus.
- **GitHub Repository:** INERIS-IA

### Family Dynamics Analysis

- Ce projet analyse les dynamiques familiales en France à l'aide de la Formal Concept Analysis (FCA).
- À partir des données du recensement national français, l'étude identifie des motifs dans les structures familiales, les foyers mono-parentaux, les décisions de fécondité et les naturalisations par mariage sur cinq zones régionales.
- Le projet explore les relations entre des variables démographiques (statut marital, composition du foyer, nationalité), offrant des éclairages sur les tendances culturelles et sociales qui façonnent les ménages français.
- **GitHub Repository:** Family Dynamics Analysis
- **Paper:** PDF

### Real-Fake Face Detection

- Présente le processus de construction et d'entraînement d'un réseau de neurones pour des tâches de classification.
- Détaille le prétraitement des données, la conception de l'architecture du modèle, l'entraînement, l'évaluation et la visualisation des résultats afin d'estimer les performances et la précision des prédictions.
- **GitHub Repository:** real-fake-detection

### EESMACF

- Évalue deux modèles NLP — SBERT (Sentence-BERT) et MiniLM — pour la classification d'analogies à partir du dataset FrameNet.

- Objectif : identifier les analogies valides et invalides en exploitant les embeddings SBERT et un classifieur MiniLM fine-tuné.
- MiniLM a atteint 99 % de précision pour distinguer les analogies valides des analogies invalides, surpassant SBERT (~55 % de précision).
- **GitHub Repository:** EESMACF
- **Paper:** PDF

## DeGatto

- Framework d'analyse de sentiments pour des avis e-commerce sur des vêtements féminins.
- Utilise un dataset Kaggle (23 000+ phrases) annoté au niveau des aspects (matière, taille, design, confort).
- Évaluation de plusieurs modèles NLP/DL/ML (BiLSTM, SVM, Logistic Regression, Multinomial Naive Bayes). BiLSTM obtient les meilleurs résultats pour l'analyse au niveau phrase, et LinearSVC performe bien pour l'analyse au niveau aspect.
- Inclut un outil de visualisation développé avec ReactJS et NodeJS pour afficher les résultats sous forme d'histogrammes ou de camemberts.
- **GitHub Repository:** DeGatto
- **Paper:** PDF

---

## Enseignement

- Assistant d'enseignement à l'IUT : *Introduction au développement logiciel* 41,5 h de TP
- Assistant d'enseignement à l'IUT : *Maintenance logicielle* 12 h de TD
- Assistant d'enseignement à Polytech : *Base de données* 16 h de TP
- Assistant d'enseignement à Polytech : *Internet* 12 h de TP

---

## Présentations

### ChatPharo: An Open Architecture for Understanding How to Talk Live to LLMs — ESUG 2025

- **Slides:** PDF
- **Video:** *Bientôt disponible*

---

## Prix et distinctions

### Prix :

- Prix du meilleur article (3e place) — IWSLT 2025, Gdańsk, Pologne  
*"Package-Aware Approach for Repository-Level Code Completion in Pharo"*

### Distinctions :

- Témoignage d'ancien étudiant mis en avant — IDMC 2024, Université de Lorraine  
*"Alumni testimonial in NLP & AI featured on the official IDMC Institute website."* [Link](#)

---

## Technologies & Compétences

- **Langages de programmation :** Python, Pharo, Java, JavaScript
- **Machine Learning & Deep Learning :** TensorFlow, PyTorch, Keras, Scikit-learn
- **Traitement Automatique des Langues (NLP) :** spaCy, NLTK
- **Big Data & systèmes distribués :** Apache Spark, Hadoop
- **Analyse de données & visualisation :** Pandas, Matplotlib, Seaborn
- **Calcul scientifique :** NumPy, SciPy
- **Développement web :** Flask, Django, Laravel, ReactJS
- **Bases de données & cloud :** MongoDB, MySQL, AWS
- **Systèmes d'exploitation :** macOS, UNIX, Windows
- **Domaines de recherche :** Intelligence artificielle, Génie logiciel, AI4SE, SE4AI

---

## Bénévolat

### NeuroTech-Lille — Président (2024–Présent)

- Pilotage du bureau exécutif et coordination des activités, projets et événements de l'association.
- Gestion du budget annuel : planification financière, suivi des dépenses et allocation des financements.

- Représentation de l'association lors de réunions avec l'université, des partenaires et des sponsors.
- Supervision de l'organisation d'événements, d'ateliers et de conférences liés aux neurotechnologies.
- Encadrement des membres de l'équipe et facilitation de la collaboration entre pôles.
- Définition d'objectifs stratégiques et garantie de l'alignement avec la mission de l'association.

#### **Ministère de la Transition Écologique — Contributeur (Club IA et Transition écologique) (2023–Présent)**

- Contribution à des projets IA soutenant des initiatives de transition écologique au sein de l'État.
- Aide au développement et à l'évaluation de cas d'usage IA pour les politiques publiques et l'impact environnemental.
- Collaboration avec des équipes pluridisciplinaires (décideurs, chercheurs, experts techniques).
- Participation à des échanges stratégiques sur un déploiement responsable et durable de l'IA dans le secteur public.

#### **Conseil National du Numérique (CNNum) — Contributeur (Café IA) (2023–Présent)**

- Contribution aux discussions nationales sur l'IA, les politiques numériques et l'impact sociétal.
- Participation à la préparation et à l'animation des sessions Café IA réunissant experts, chercheurs et parties prenantes.
- Participation à l'analyse de tendances IA, de cas d'usage et d'enjeux réglementaires.
- Appui aux initiatives de partage de connaissances autour d'une IA responsable et éthique.

#### **Croix-Rouge Française — Secouriste bénévole (2022–2023)**

- Réalisation de gestes de premiers secours et assistance d'urgence lors d'événements et d'opérations.
- Prise en charge d'urgences médicales selon les protocoles de sécurité et d'intervention.
- Collaboration avec des équipes pluridisciplinaires en contexte de forte pression.
- Maintien des compétences via des formations régulières en secourisme et réponse d'urgence.

---

### **Loisirs**

- **Activités nautiques** : Natation · Plongée
- **Sports de plein air et d'aventure** : Randonnée · Ski · Saut en parachute
- **Activités créatives et artistiques** : Photographie
- **Sports de précision** : Boxe · Bowling · Golf

---

### **Langues**

- **Arabe** : Langue maternelle
- **Français** : Langue maternelle
- **Anglais** : B2

*Dernière mise à jour : February 22, 2026*