# Omar ABEDELKADER

⚲ Lille, France    ✉ omar.abedelkader@inria.fr    🔗 omarabedelkader.fr

## About Me

I am an AI researcher with extensive experience in Artificial Intelligence and Software Engineering. Currently pursuing a Ph.D. in Computer Science at the University of Lille.

## Education

**University of Lille** *Villeneuve-d'Ascq, Nord*      *Oct. 2024 – Oct. 2027*
*Ph.D in Computer Science*

**University of Lorraine** *Nancy, M&M*      *Sept. 2022 – Sept. 2024*
*M.Sc. in Natural Language Processing*

**Lebanese University** *Beirut, Lebanon*      *Sept. 2019 – July 2022*
*B.Sc. in Data Science*

## Experience

**AI Researcher**      *Villeneuve d'Asq, Nord*
*INRIA*      *Oct. 2024 – Oct. 2027*

- Improving code completion and generation using LLMs, specifically targeting the Pharo programming language, which has limited training data.
- Developing techniques for code completion, type inference, and deployment in Pharo's IDE, with a focus on runtime performance.

**AI Engineer**      *Verneuil-en-Halatte, Oise*
*INERIS*      *Sept. 2023 – Sept. 2024*

- Developing "INERIS-IA," a tool to classify textual documents based on INERIS's strategic goals using ML and NLP techniques
- Create boolean queries for document retrieval, and improve corpus quality through document similarity and keyword extraction.

**Intern - AI Researcher**      *Villetaneuse, SSD*
*LIPN*      *June 2023 – Aug. 2023*

- Comparison of various sampling techniques for probabilistic planning, particularly in generating literary narratives.
- Evaluate different methods, including the Score Function Estimator (SFE) and more advanced techniques like Gumbel-Softmax, to assess their effectiveness in creating coherent and creative stories.

## Projets

**Family Dynamics Analysis**      *Nancy, M&M*
*IDMC*      *Sept. 2023 – Jan. 2024*

- This project analyzes family dynamics in France using Formal Concept Analysis (FCA). By leveraging data from the French national census, the study identifies patterns in family structures, single-parent households, childbearing decisions, and naturalization through marriage across five regional zones. Using advanced data mining techniques, the project explores relationships between demographic variables such as marital status, household composition, and nationality, offering insights into the cultural and social trends shaping French households.
- **GitHub Repository**: [GitHub ↗](#)
- **Paper**: [Paper ↗](#)

**Real-Fake Face detection**      *Nancy, M&M*
*IDMC*      *Sept. 2023 – Jan. 2024*

- This project demonstrates the process of building and training a neural network for classification tasks using machine learning frameworks. It walks through data preprocessing, model architecture definition, training, evaluation, and visualization of results, providing insights into model performance and prediction accuracy.
- **GitHub Repository**: GitHub↗

**EESMACF** — *Nancy, M&M*
*IDMC & LORIA* — *Sept. 2022 – June. 2023*

- This project focuses on evaluating the effectiveness of two NLP models—SBERT (Sentence-BERT) and MiniLM—in classifying analogies using the FrameNet dataset. FrameNet is a rich semantic database that captures the meanings of lexical units within specific contexts called frames. The objective is to identify valid and invalid analogies in FrameNet, leveraging the capabilities of SBERT and a fine-tuned MiniLM model. SBERT is used to create dense vector embeddings of sentences that retain semantic similarity, while MiniLM is trained and fine-tuned to classify analogies effectively. The project found that MiniLM achieved an impressive 99% accuracy in distinguishing between valid and invalid analogies, outperforming SBERT, which had around 55% accuracy.
- **GitHub Repository**: GitHub↗
- **Paper**: Paper↗

**DeGatto** — *IDMC, M&M*
*IDMC* — *Sept. 2022 – Jan. 2023*

- The "DeGatto" project is a sentiment analysis framework designed for e-commerce, specifically focusing on women's apparel reviews. Using a dataset sourced from Kaggle, which includes over 23,000 sentences and aspect-level annotations for material, size, design, and comfort, the project aims to support e-commerce businesses and customers by analyzing feedback at both the sentence and aspect levels. Various NLP, DL, and ML models were tested, including LSTM (BiLSTM), SVM, Logistic Regression, and Multinomial Naive Bayes, with BiLSTM achieving the best results in sentence-level analysis and LinearSVC performing well at aspect-level analysis. A visualization tool was developed using ReactJS and NodeJS, enabling users to view results as bar or pie charts. The project concluded that BiLSTM is the most suitable model for sentence-level sentiment analysis, while LinearSVC excels in aspect-level analysis, providing a robust framework for sentiment classification in e-commerce contexts.
- **GitHub Repository**: GitHub↗
- **Paper**: Paper↗

## Technologies & Skills

**Programming Languages:** Python, Pharo, Java, R, JavaScript, PHP, MySQL

**Machine Learning & Deep Learning:** TensorFlow, Keras, Scikit-learn

**Natural Language Processing (NLP):** SpaCy, NLTK

**Big Data & Distributed Systems:** Apache Spark, Hadoop

**Data Analysis & Visualization:** Pandas, NumPy, Matplotlib, Seaborn

**Frameworks & Tools:** Flask, Django, MongoDB, Laravel, ReactJS, AWS, PyCharm,

**Operating Systems:** MacOS, UNIX, Windows

**Scientific Computing:** SciPy, NumPy

**Research Areas:** Artificial Intelligence, Software Engineering

## Volunteering

**NeuroTech-Lille:** President

**Ministère de la Transition Écologique:** Contributor in Club IA et Transition écologique

**Conseil National du Numérique (CNNum):** Contributor in Café IA

**Croix-Rouge Française:** Volunteer First Aider

## Hobbies

**Water Activities:** Swimming · Diving

**Outdoor and Adventure Sports:** Climbing · Skiing · Skydiving

**Creative and Artistic Activities:** Photography

**Precision Sports:** Boxing · Bowling · Golf

## Language

**French:** Native

**English:** B2

**Scan Me !**