# Ilias Azizi


30 March 1996

azizii.ilias@gmail.com

+33 18 10 31 50

LinkedIn/Ilias-Azizi

9B, Bd Jourdan, 75014 - Paris, France


## Certificate

Deep Learning Specialization
Organization: Deeplearning.ai

Data Scientist
Organization: DataCamp

Data Analyst
Organization: DataCamp

## Languages

Tamazight: Native
Arabic: Native (Bilingual)
French: Fluent
English: Fluent

## Skills

Research Skills: Information Retrieval, Deep Learning, Recommendation Systems, NLP/LLM, Image Processing, Distributed Systems, Algorithm Optimization, Literature Review, Research Methodology
Programming Languages: C++, Python, Scala, Rust, Java
DBMS: Sql Server, Oracle, MongoDB
Big Data Technologies: Hadoop, Cassandra, Kafka, Spark, MapReduce

## Research Interests

My primary research area revolves around scalable similarity search, information retrieval and machine learning applied to large, high-dimensional datasets.

## Education

2020 - 2024 — PhD Student: Scalable Machine Learning on Massive High-Dimensional Vectors — Benguerir, Morocco - Paris, France
Supervisors: Prof. Karima Echihabi (Mohammed VI Polytechnic University) and Prof. Themis Palpanas (Université Paris-Cité)

2019 - 2020 — Pre-Doctoral Program DNA: Data Science, Networking, and Algorithm Thinking — Benguerir, Morocco
Mohammed VI Polytechnic University

2017 - 2019 — Masters in Data Science and Big Data — Rabat, Morocco
High National School Of Computer Science And Systems Analysis (ENSIAS) - Mohammed V University
Honors: First of the cohort with highest honors

2014 - 2017 — B.Sc Mathematics and Computer Science — Agadir, Morocco
Faculty of Science - University of Ibn Zohr

## Academic and Professional Experience

2024 — Applied Scientist Intern — Amazon.com Services LLC, USA
Join the Alexa AI team at Amazon for a three months internship as an Applied Scientist Intern starting in September 2024, where I will design and implement solutions for vector search engines within the Amazon ecosystem to enhance the scalability and efficiency of Amazon's information retrieval operations.

2023 — Undergraduate Project Supervisor — Mohammed VI Polytechnic University, Morocco
Provided mentorship for undergraduate projects, focusing on graph-based search strategies, ensuring timely progress, and fostering excellence in research methodology and problem-solving.

2020 - 2024 — External Reviewer, VLDB and SIGMOD Conferences
Reviewed submissions for VLDB and SIGMOD conferences, offering detailed evaluations and improvement suggestions to enhance paper quality.

2022 - 2023 — Teaching Assistant — Université Paris-Cité, France
Prepared and delivered labs on programming in Python for undergraduate students.

2022 - 2023 — Teaching Assistant — Mohammed VI Polytechnic University, Morocco
Assisted undergraduate students in lectures and labs on Database Management.

2022 — Teacher Assistant - UM6P AI for Science Bootcamp — Mohammed VI Polytechnic University, Morocco
Assisted researchers and developers in applying AI tools during NVIDIA and UM6P GPU Bootcamp.

2019 — Master's Thesis : Strong Tail Recommendation System — ENSIAS
Proposed a noval metric, strong tail similarity, to model the relation between items in a bipartite graph of user-item. Developed a hybrid recommender system framework for generating ranked recommendation lists based on the proposed strong tail similarity, graph embeddings, community detection, and a deep neural network. Technologies used: Node2vec, Deep Neural Network, Pytorch, Python

## Invited Talks and Presentations

2024 — 50th International Conference on Very Large Data Bases — Guangzhou, China
Presented our published Phd Workerhop paper at 50th International Conference on Very Large Data Bases 2024.

2024 — FAISS Team - FAIR (Meta) Seminar — Paris, France
Delivered a talk on graph-based similarity search and our approach ELPIS for similarity search over high-dimensional large datasets.

2023 — College of Computing Seminar - UM6P — Benguerir, Morocco
Presented results and work on similarity search and graph-based approaches at the prestigious conference VLDB 2023.

2023 — 49th International Conference on Very Large Data Bases — Vancouver, Canada
Presented our published work at 49th International Conference on Very Large Data Bases 2023.

2022 - 2023 — DiNo Research Group Seminar — Paris, Francce
Delivered multiple presentations on large-scale graph-based similarity search using high-dimensional vectors, focusing on potential publications, at the DiNo research seminar on Data Analytics, Université Paris-Cité.

## Publications

2024 — Ilias Azizi, Karima Echihabi, and Themis Palpanas. "Graph-Based Vector Search: An Experimental Evaluation of the State-of-the-Art". Proceedings of the ACM on Management of Data, 2025
We present a detailed survey of the state-of-the-art graph-based methods for in-memory ng-approximate similarity search, with a new taxonomy based on construction approaches with an evaluation of key design choices for efficient graph-based similarity search at large scale datasets of high dimensional vectors.

2024 — Ilias Azizi. "Vector Search on Billion-Scale Data Collections". VLDB 2024 Ph.D. Workshop.
The Workshop paper presented at the VLDB 2024 PhD workshop addresses the problem of large-scale vector search by introducing ELPIS, an in-memory vector search technique that blends the strengths of graph and tree-based index structures while addressing their limitations. The paper also provides important insights into graph structures for similarity search and includes findings from an in-depth investigation of various components of graph-based methods.

2022 — Ilias Azizi, Karima Echihabi, and Themis Palpanas. "Elpis: Graph-based similarity search for scalable data science". Proceedings of the VLDB Endowment, 2023, vol. 16, no 6, p. 1548-1559.
49th International Conference on Very Large Data Bases, Vancouver, Canada.
We proposed a indexing approach, ELPIS, for similarity search on large scale datasets of high-dimensional vectors. ELPIS builds an index 2-8x faster than the best competitors up to 40% less memory footprint, and achieves a high recall of 0.99, up to 2x faster than state-of-the-art methods. We also shared valuable insights and lessons learned including the key design choices that help optimize graph-based similarity search performance by adopting sophisticated divide-and-conquer design for both indexing and search.

## Grants and Awards

2024 — VLDB Travel Grant
Covered expenses for attending the 50th International Conference on Very Large Data Bases at Guangzhou,China

2023 — VLDB Travel Grant
Covered expenses for attending the 49th International Conference on Very Large Data Bases at Vancouver,Canada