

Ali Maatouk

Researcher on Foundation Models for Structured Data

📍 New York Metropolitan Area | ✉ ali.maatouk@yale.edu | 🔗 ali-maatouk.github.io

Summary

I am a researcher at the Yale Institute for Foundations of Data Science, where I work on foundation models for structured data. Our work focuses on multimodal learning methods that combine geometric and graph-based approaches to analyze text, time-series data, and graph structures, with applications in observability domains and finance. I have published over 40 papers in the top venues of the fields of artificial intelligence, signal processing, and information theory, and have received multiple research awards for my work.

Professional Experience

Yale Institute for Foundations of Data Science

New Haven, CT

Researcher

Jan 2024 - Present

- Developed geometry-aware foundation models, including non-Euclidean large language models using a mixture-of-space-experts architecture
- Designed graph-based learning methods that integrate graph representations into large language models for enhanced retrieval and reasoning
- Engaged in industry-sponsored research on hyperbolic and graph-augmented retrieval in collaboration with partners such as Goldman Sachs
- Mentored 10+ graduate and undergraduate students across interdisciplinary research projects

Huawei Technologies

Paris, France

Researcher

Sep 2021 - Dec 2023

- Designed and deployed machine learning models on large-scale datasets to improve network performance
- Developed optimization frameworks integrating machine learning models to enhance network efficiency and reliability
- Conducted research on the application of large language models and retrieval-augmented generation for large-scale networked systems
- Mentored interns from leading institutions, such as École Polytechnique and University of Oxford

CentraleSupélec & University of Maryland, College Park

Paris, France & College Park, MD

Researcher — Joint PhD Program in Applied Mathematics

Oct 2017 - Aug 2021

- Introduced a new performance metric for quantifying information decay in dynamical systems that has been widely adopted in the information theory community
- Developed theoretical frameworks based on stochastic hybrid systems for controlling dynamical systems to minimize information decay
- Designed Markov decision process and Restless Multi-Armed Bandit algorithms to optimize multi-user dynamical system performance

Education

CentraleSupélec & University of Maryland, College Park

Paris, France & College Park, MD

PhD in Applied Mathematics

Oct 2017 - Dec 2020

Focus: Stochastic hybrid systems and Markov decision processes for minimization of information decay in dynamical systems

CentraleSupélec

Paris, France

MS in Applied Mathematics | Ranked 1st in cohort

Sep 2016 - Sep 2017

Focus: Information theory and signal processing

Lebanese University

Beirut, Lebanon

BS in Computer Engineering | Ranked 1st in cohort

Sep 2012 - Aug 2016

Technical Skills

Programming Languages: Python, MATLAB

Deep Learning Frameworks: PyTorch, Hugging Face Transformers, DeepSpeed

Machine Learning Libraries: Scikit-learn, XGBoost, LightGBM, NumPy, SciPy, pandas

Awards and Honors

Oral Presentation at Learning on Graphs (LoG) Conference	Dec 2025
ACM KDD Blue Sky Ideas Award for our work on foundation models for public health	Aug 2025
Robert Innovation Fellowship from Yale University	Feb 2025
AI Innovation Seed Funding from Yale University	Nov 2024
Huawei Future Star Award for outstanding research contributions	Dec 2021
Plateau de Saclay PhD Award for one of the best scientific contributions	Dec 2020
André Blanc-Lapierre Award, National Prize for best Master's student in France in the domain of electricity, information technology, and communication	Nov 2017
André Blanc-Lapierre Award, Trophée Ile-de-France for best Master's project in Paris region	Nov 2017
Digicosme Scholarship for academic excellence	Jul 2016
First Place in the Lebanese Baccalauréat official exams in Beirut	Jul 2012

Publications

Google Scholar: 1,852 citations | h-index: 20 | i10-index: 27

2026

LitBench: A Graph-Centric Large Language Model Benchmarking Tool For Literature Tasks

A. Varvarigos, [Ali Maatouk](#), J. Zhang, N. Bui, J. Chen, L. Tassiulas, R. Ying
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2026

2025

Position: Beyond Euclidean-Foundation Models Should Embrace Non-Euclidean Geometries

N. He, J. Liu, B. Zhang, N. Bui, [Ali Maatouk](#), M. Yang, I. King, M. Weber, R. Ying
Learning on Graphs Conference (LoG), 2025 **(Oral Presentation)**

TRACE: Grounding Time Series in Context for Multimodal Embedding and Retrieval

J. Chen, Z. Zhao, G. Nurbek, A. Feng, [Ali Maatouk](#), L. Tassiulas, Y. Gao, R. Ying
Conference on Neural Information Processing Systems (NeurIPS), 2025

HELM: Hyperbolic Large Language Models via Mixture-of-Curvature Experts

N. He, R. Anand, H. Madhu, [Ali Maatouk](#), S. Krishnaswamy, L. Tassiulas, M. Yang, R. Ying
Conference on Neural Information Processing Systems (NeurIPS), 2025

LitFM: A retrieval augmented structure-aware foundation model for citation graphs

J. Zhang, [Ali Maatouk](#), J. Chen, N. Bui, Q. Xie, L. Tassiulas, H. Xu, J. Shao, R. Ying
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), 2025

TelecomTS: A Multi-Modal Observability Dataset for Time Series and Language Analysis

A. Feng, A. Varvarigos, I. Panitsas, D. Fernandez, J. Wei, Y. Guo, J. Chen, [Ali Maatouk](#), L. Tassiulas, R. Ying
arXiv preprint arXiv:2510.06063, 2025

MTBench: A Multimodal Time Series Benchmark for Temporal Reasoning and Question Answering

J. Chen, A. Feng, Z. Zhao, J. Garza, G. Nurbek, C. Qin, [Ali Maatouk](#), L. Tassiulas, Y. Gao, R. Ying
arXiv preprint arXiv:2503.16858, 2025

Age of Information Optimization with Preemption Strategies for Correlated Systems

E. Erbayat, [Ali Maatouk](#), P. Zou, S. Subramaniam
IEEE International Symposium on Information Theory (ISIT), 2025

Interdisciplinary development and fine-tuning of cardio, a large language model for cardiovascular health education in HIV care: Tutorial

R. Rullo, [Ali Maatouk](#), T. Huang, J. Chen, W. Qiu, G. O'Connor, J. Womack, T. Sadak, C. Rodriguez, P. Carneiro, T. Espinosa, A. Marshall, R. Ying, S. Ramos
Journal of Medical Internet Research, Vol. 27, e77053, 2025

Telco-oRAG: Optimizing Retrieval-augmented Generation for Telecom Queries via Hybrid Retrieval and Neural Routing

A. L. Bornea, F. Ayed, A. De Domenico, N. Piovesan, T. S. Salem, [Ali Maatouk](#)
IEEE Journal on Selected Areas in Communications, 2025

TeleQnA: A benchmark dataset to assess large language models telecommunications knowledge

[Ali Maatouk](#), F. Ayed, N. Piovesan, A. De Domenico, M. Debbah, Z. Q. Luo
IEEE Network, 2025

Agoran: An Agentic Open Marketplace for 6G RAN Automation

I. Chatzistefanidis, N. Nikaein, A. Leone, [Ali Maatouk](#), L. Tassiulas, R. Morabito, I. Pitsiorlas, M. Kountouris
Computer Networks, 111927, 2025

Penalty Upon Decision: A Metric to Quantify Decision Costs in Status Update Systems

P. Zou, [Ali Maatouk](#), J. Zhang, S. Subramaniam
IEEE Transactions on Networking, 2025

SANDWICH: Towards an offline, differentiable, fully-trainable wireless neural ray-tracing surrogate

Y. Jin, [Ali Maatouk](#), S. Girdzijauskas, S. Xu, L. Tassiulas, R. Ying
IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN), 2025

Large-Scale AI in Telecom: Charting the Roadmap for Innovation, Scalability, and Enhanced Digital Experiences

A. Shahid, A. Kliks, A. Al-Tahmeesschi, A. Elbakary, A. Nikou, [Ali Maatouk](#), et al.
arXiv preprint arXiv:2503.04184, 2025

2024

Pay Attention to What Matters

P. L. Silva, A. de Domenico, [Ali Maatouk](#), F. Ayed
MINT@NeurIPS 2024, 2024

Tele-LLMs: A Series of Specialized Large Language Models for Telecommunications

[Ali Maatouk](#), K. C. Ampudia, R. Ying, L. Tassiulas
arXiv preprint arXiv:2409.05314, 2024

Age of Information Optimization and State Error Analysis for Correlated Multi-Process Multi-Sensor Systems

E. Erbayat, [Ali Maatouk](#), P. Zou, S. Subramaniam

International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc), 2024

Large language models for telecom: Forthcoming impact on the industry

Ali Maatouk, N. Piovesan, F. Ayed, A. De Domenico, M. Debbah
IEEE Communications Magazine, 2024

Telco-RAG: Navigating the Challenges of Retrieval Augmented Language Models for Telecommunications

A. L. Bornea, F. Ayed, A. De Domenico, N. Piovesan, Ali Maatouk
IEEE Global Communications Conference (GLOBECOM), 2024

Hermes: A Large Language Model Framework on the Journey to Autonomous Networks

F. Ayed, Ali Maatouk, N. Piovesan, A. De Domenico, M. Debbah, Z. Q. Luo
arXiv preprint arXiv:2411.06490, 2024

Predictive handover strategy in 6G and beyond: A deep and transfer learning approach

I. Panitsas, A. Mudvari, Ali Maatouk, L. Tassiulas
arXiv preprint arXiv:2404.08113, 2024

A framework for the evaluation of network reliability under periodic demand

Ali Maatouk, F. Ayed, S. Biao, W. Li, H. Bao, E. Zio
IEEE/ACM Transactions on Networking, Vol. 32(3), 2495-2510, 2024

2023

FlexTrain: A Dynamic Training Framework for Heterogeneous Devices Environments

M. Unsal, Ali Maatouk, A. De Domenico, N. Piovesan, F. Ayed
WANT@NeurIPS 2023, 2023

Age-aware stochastic hybrid systems: Stability, solutions, and applications

Ali Maatouk, M. Assaad, A. Ephremides
IEEE Journal on Selected Areas in Information Theory, Vol. 4, 762-783, 2023

An Optimization Framework For Anomaly Detection Scores Refinement With Side Information

Ali Maatouk, F. Ayed, W. Li, Y. Wang, H. Zhu, J. Ye
IEEE Global Communications Conference (GLOBECOM), 6043-6049, 2023

How Costly Was That (In) Decision?

P. Zou, Ali Maatouk, J. Zhang, S. Subramaniam
International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), 2023

2022

Analysis of an age-dependent stochastic hybrid system

Ali Maatouk, M. Assaad, A. Ephremides
IEEE International Symposium on Information Theory (ISIT), 150-155, 2022

The age of incorrect information: An enabler of semantics-empowered communication

Ali Maatouk, M. Assaad, A. Ephremides
IEEE Transactions on Wireless Communications, Vol. 22(4), 2621-2635, 2022

Semantics-empowered communications through the age of incorrect information

Ali Maatouk, M. Assaad, A. Ephremides

IEEE International Conference on Communications (ICC), 3995-4000, 2022

Timely updates with priorities: Lexicographic age optimality

Ali Maatouk, Y. Sun, A. Ephremides, M. Assaad

IEEE Transactions on Communications, Vol. 70(5), 3020-3033, 2022

2021

On the Global Optimality of Whittle's Index Policy for Minimizing the Age of Information

S. Kriouile, M. Assaad, Ali Maatouk

IEEE Transactions on Information Theory, Vol. 68(1), 572-600, 2021

2020

Asymptotically optimal scheduling policy for minimizing the age of information

Ali Maatouk, S. Kriouile, M. Assaad, A. Ephremides

IEEE International Symposium on Information Theory (ISIT), 1747-1752, 2020

The age of incorrect information: A new performance metric for status updates

Ali Maatouk, S. Kriouile, M. Assaad, A. Ephremides

IEEE/ACM Transactions on Networking, Vol. 28(5), 2215-2228, 2020

On the age of information in a CSMA environment

Ali Maatouk, M. Assaad, A. Ephremides

IEEE/ACM Transactions on Networking, Vol. 28(2), 818-831, 2020

On the optimality of the Whittle's index policy for minimizing the age of information

Ali Maatouk, S. Kriouile, M. Assaad, A. Ephremides

IEEE Transactions on Wireless Communications, Vol. 20(2), 1263-1277, 2020

Status updates with priorities: Lexicographic optimality

Ali Maatouk, Y. Sun, A. Ephremides, M. Assaad

International Symposium on Modeling and Optimization in Mobile, Ad hoc, and Wireless Networks (WiOpt), 2020

2019

Age of information with prioritized streams: When to buffer preempted packets?

Ali Maatouk, M. Assaad, A. Ephremides

IEEE International Symposium on Information Theory (ISIT), 325-329, 2019

Energy efficient and throughput optimal CSMA scheme

Ali Maatouk, M. Assaad, A. Ephremides

IEEE/ACM Transactions on Networking, Vol. 27(1), 316-329, 2019

Minimizing the age of information in a CSMA environment

Ali Maatouk, M. Assaad, A. Ephremides

International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2019

Minimizing the age of information: NOMA or OMA?

Ali Maatouk, M. Assaad, A. Ephremides

IEEE INFOCOM Conference on Computer Communications Workshops, 2019

2018

The age of updates in a simple relay network

Ali Maatouk, M. Assaad, A. Ephremides

IEEE Information Theory Workshop (ITW), 1-5, 2018

On optimal scheduling for joint spatial division and multiplexing approach in FDD massive MIMO

Ali Maatouk, S. E. Hajri, M. Assaad, H. Sari

IEEE Transactions on Signal Processing, Vol. 67(4), 1006-1021, 2018

Stay Longer at the Network's Edge: A Novel Proactive Caching Policy through Sojourn Time

J. Denis, Ali Maatouk, S. E. Hajri, M. Assaad

IEEE Global Communications Conference (GLOBECOM), 1-6, 2018

A simple NOMA scheme with optimum detection

E. Caliskan, Ali Maatouk, M. Koca, M. Assaad, G. Gui, H. Sari

IEEE Global Communications Conference (GLOBECOM), 1-6, 2018

Graph theory based approach to users grouping and downlink scheduling in FDD massive MIMO

Ali Maatouk, S. E. Hajri, M. Assaad, H. Sari, S. Sezginer

IEEE International Conference on Communications (ICC), 1-7, 2018

On the foundation of NOMA and its application to 5G cellular networks

H. Sari, Ali Maatouk, E. Caliskan, M. Assaad, M. Koca, G. Gui

IEEE Wireless Communications and Networking Conference (WCNC), 1-6, 2018

Frequency-domain NOMA with two sets of orthogonal signal waveforms

Ali Maatouk, E. Çalışkan, M. Koca, M. Assaad, G. Gui, H. Sari

IEEE Communications Letters, Vol. 22(5), 906-909, 2018

Invited Talks

Global System for Mobile Communications Association | London, UK | September 2025

Multi-Modal Data Integration for Telecom AI Models

Karlsruhe Institute of Technology | Karlsruhe, Germany | July 2023

Goal-Oriented Semantics Communications: The Role of Data Freshness

Massachusetts Institute of Technology | Cambridge, MA | February 2023

Large Scale Optimization of Cellular Networks: The Role of Machine Learning

Yale University | New Haven, CT | February 2023

Large Scale Optimization of Cellular Networks: The Role of Machine Learning

George Washington University | Washington DC | November 2022

Semantics-Empowered Communications: A New Paradigm for Networked Systems

Middle East Technical University | Ankara, Turkey | April 2021

Toward a Semantics-Empowered Communication Paradigm

École Nationale Supérieure de l'Électronique et de ses Applications | Cergy, France | March 2021

Semantics of Data: Age of Information and Beyond

Auburn University | Auburn, AL | March 2020

The Age of Incorrect Information: A New Perspective on Status Update Systems

Tutorials

IEEE Global Communications Conference (GLOBECOM) | Madrid, Spain | December 2021

Age of Information: A Paradigm Shift Towards Semantic Properties in Data Science

IEEE International Conference on Communications (ICC) | Montreal, Canada | June 2021

Semantics in Communications: The Example of Information Freshness

Academic Services

AI/ML Community

- Reviewer for NeurIPS, ICML, ICLR, and KDD

Information Theory Community

- TPC member for IEEE ISIT and IEEE INFOCOM Workshop on Age and Semantics of Information
- Reviewer for IEEE ISIT, IEEE ITW, IEEE Transactions on Information Theory, and IEEE Journal on Selected Areas in Information Theory

Signal Processing Community

- Organizer of the Age of Information and Beyond symposium at the 57th Annual Asilomar Conference
- Reviewer for the Asilomar Conference and IEEE Transactions on Signal Processing

Networking Community

- TPC member for netAIsys @ ACM MobiSys, IEEE ICC Workshop on The Impact of LLMs on 6G Networks, and IEEE GLOBECOM IMMLLM6G Workshop
- Reviewer for IEEE Infocom, IEEE/ACM Transactions on Networking, and IEEE Journal on Selected Areas in Communications

Mentorship & Supervision

Yale University

- **PhD Students:** Jiasheng Zhang, Hiren Madhu, Trey Crump, Andreas Varvarigos, Ioannis Panitsas, Iason Ofeidis, Eftychia Makri
- **Undergraduate Students:** Neil He, Rishabh Anand, Austin Feng, Rawad Kansoh

George Washington University

- **PhD Students:** Egemen Erbayat, Peng Zhou

University of Maryland, College Park

- **PhD Students:** Yutao Chen, Zahra Zare

Huawei Technologies

- **Interns:** Mert Unsal, Andrei-Laurentiu Bornea, Pedro Luiz Silva, Shi Biao, Henrique Aguiar

Languages

English: Fluent

French: Native

Arabic: Native

Spanish: Elementary