DR. MERYEM M'HAMDI

RESEARCH SCIENTIST / MACHINE LEARNING & NLP

✉ meryemmhamdi1@gmail.com ⌖ Redmond, WA ◇ meryemmhamdi1.github.io

linkedin.com/in/meryem-m-hamdi-phd-69964980/ scholar.google.com/citations?user=WGiywe8AAAAJ

PhD-trained Machine Learning Researcher with expertise in multilingual NLP, continual learning, meta-learning, and data-efficient adaptation. First-author publications at ACL, NAACL, and SIGIR, with research and industry experience spanning USC, Adobe Research, Meta, and Microsoft. Combines foundational machine learning research with applied expertise in LLM alignment, post-training evaluation, and production-scale search and retrieval systems. Highly adaptable to multicultural and challenging environments, collaborative, and driven by intellectual curiosity, with a passion for developing reliable, efficient, and generalizable AI systems that bridge research innovation and real-world impact.

PROFESSIONAL & RESEARCH EXPERIENCE

Meta | Research Collaborator V | Dec 2024 – Present

Generative AI research and evaluation • via Rose International • Remote, Redmond, WA

- Led golden data annotation, quality benchmarks, and company continuous dogfooding across coding, STEM, finance, and medicine, coordinating more than 10 annotators.
- Contributed to LLM evaluation and quality improvement initiatives spanning response grounding, relevance, formatting, and location-based user experiences, supporting an increase in chip and map engagement from 14% to 19% and local rendering coverage from 20% to 83% of local responses.
- Led the development and three-stage QA of 296 coding-agent evaluation tasks over three months, yielding 53 high-quality finalized tasks with 24.9 grading rubrics per task, 97% LLM–human grading agreement, and over 91% model differentiation at a 30% threshold, outperforming external vendor benchmarks.
- Developed and validated 5,000+ evaluation rubrics and annotations supporting LLM post-training, alignment, and reasoning assessment across technical domains.

Microsoft AI | Data Scientist II | Mar 2024 – Aug 2024

Bing Multimedia • Redmond, WA

- Developed FAIRY-Search, an LLM-based framework for evaluating personalized image-search relevance across user profiles and experimental flights.
- Designed comparative evaluation criteria and offline experiments to analyze personalized ranking quality and model behavior.

USC Information Sciences Institute | Research Assistant | Jan 2019 – Jan 2024

CUTELABNAME Research Group (Prof. Jonathan May) • Marina del Rey, CA

- Introduced cross-lingual continual learning benchmarks and evaluation protocols for knowledge preservation, accumulation, and generalization.
- Investigated meta-learning and cross-lingual transfer for low-resource language adaptation, event extraction, and task-oriented NLP.
- Proposed Leitner-guided memory replay to reduce catastrophic forgetting while balancing plasticity and stability.

RESEARCH INTERNSHIPS

Microsoft | Data & Applied Scientist PhD Intern | May 2023 – Aug 2023

Bing Multimedia / WebXT • Redmond, WA

- Researched and implemented a dual-tower deep ranking approach with triplet-loss learning for multimedia search relevance.
- Validated ranking improvements contributing to a reported 12% increase in user engagement in production.

Adobe Research | Research Scientist Intern | May 2022 – Nov 2022

Natural Language Processing Group • San Jose, CA

- Proposed MAML-Align, a multilingual meta-distillation approach to semantic retrieval, evaluated on low-resource and unseen-language transfer (SIGIR 2024).
- Designed multilingual few-shot adaptation and embedding experiments using LAREQA and STSB-Multi.

RESEARCH EXPERTISE

Learning & Adaptation

Cross-lingual transfer, continual learning, meta-learning, low-resource NLP, knowledge distillation

Generative AI & Evaluation

LLM alignment, SFT/RLHF evaluation, reasoning benchmarks, prompt optimization, rubric design

Retrieval & Representation

Multilingual embeddings, semantic search, deep ranking, dual-tower models, triplet loss

Programming & Frameworks

Python, PyTorch, TensorFlow, SQL, Java, C++, Git

EDUCATION

Ph.D. in Computer Science

USC, USA

- Research Interests: NLP, Multilinguality, Information Extraction, Meta-Learning, Continual Learning, Human Learning.
- GPA: 3.92/4.
- Thesis Title: Towards More Human-Like Cross-Lingual Transfer Learning.

M.Sc. in Computer Science

EPFL, Switzerland

- Specialization: Data Analytics.
- GPA: 5.29/6.
- Thesis title: Creation and evaluation of multilingual embeddings for downstream semantic tasks.

Honors B.Sc. in Computer Science

Al Akhawayn University in Ifrane, Morocco

- Minors: Mathematics & Business Administration.
- GPA: 3.91/4.
- Leadership Development Institute.
- Exchange program at GWU.

SELECTED PUBLICATIONS

- Multilingual Meta-Distillation Alignment for Semantic Retrieval. **SIGIR 2024**.
- Leitner-Guided Memory Replay for Cross-lingual Continual Learning. **NAACL 2024**.
- Cross-Lingual Continual Learning. **ACL 2023**.
- X-METRA-ADA: Cross-lingual Meta-Transfer Learning Adaptation to Natural Language Understanding. **NAACL 2021**.

- Contributed to work associated with US patent publication US12547901B2.

Adobe Research | NLP Research Intern | May 2020 – Nov 2020

Natural Language Processing Group • Remote

- Investigated meta-transfer learning for rapid cross-lingual adaptation in NLU and question answering (NAACL 2021).
- Contributed to work associated with US patent publication US20230297603A1.

Swisscom / EPFL | Data Science Research Intern | Feb 2018 – Sep 2018

Multilingual embeddings • Lausanne, Switzerland

- Developed a multilingual embedding evaluation framework spanning classification, event detection, and churn-related tasks (CoNLL 2018).
- Explored multilingual representations for unsupervised event detection across 11 languages.

SELECTED TECHNICAL & RESEARCH PROJECTS

Cross-Lingual Continual Learning | USC / ISI

• Developed a benchmark for sequential multilingual adaptation, measuring knowledge preservation, accumulation, and generalization (ACL 2023)

• Designed a spaced-repetition-inspired replay strategy to mitigate catastrophic forgetting and study plasticity–stability trade-offs (NAACL 2024).

Cross-lingual Meta-Transfer Learning | Adobe Research

• X-METRA-ADA: Cross-Lingual NLU: Investigated MAML-based rapid adaptation for task-oriented dialogue and question answering across languages (NAACL 2021).

• MAML-Align: Multilingual Semantic Retrieval: Developed multilingual meta-distillation for low-resource semantic retrieval, evaluating few-shot transfer on LAREQA and STSB-Multi (SIGIR 2024).

Cross-Lingual Information Extraction | USC / ISI

• Implemented contextualized cross-lingual event trigger extraction using multilingual BERT and transfer learning; evaluated performance against linguistic baselines (CoNLL 2019).

• Evaluated language-independent transfer against linguistic baselines for event trigger detection in low-resource settings.

Multilingual Embedding Evaluation | EPFL / Swisscom

• Built an end-to-end evaluation library for multilingual representations across text classification, event detection and churn-related applications.

• Compared multilingual representations on downstream semantic tasks and unsupervised event detection across 11 languages.

NLP Processing Pipeline | EPFL

• Implemented lexical analysis with finite-state automata, morphological processing with transducers and CYK syntactic parsing for semantic similarity research.

• Built a modular linguistic pre-processing pipeline combining lexical, morphological and syntactic analysis for semantic similarity experiments.

Emotion Analytics & Visualization | EPFL

• Developed a platform for analyzing approximately 12 million geolocated tweets and visualizing regional emotions with D3.js.

• Integrated fine-grained emotion analysis with an interactive D3.js visualization of canton-level trends.

- Contextualized Cross-Lingual Event Trigger Extraction with Minimal Resources. **CoNLL 2019**.
- Churn Intent Detection in Multilingual Chatbot Conversations and Social Media. **CoNLL 2018**.

LEADERSHIP & SERVICE

Conference Leadership | WiNLP Organizing Committee

May 2025–Present

- WiNLP chair organizer and area chair

Peer Review | Top-Tier NLP Conferences

May 2020 – Present

- EMNLP '21–'24; ACL '21, '23–'24; EACL '23; NAACL '22, '24; AACL-IJCNLP '20, '22; MetaNLP '21; ARR '22–'24.

Academic Service | Natural Language Seminar Organizer

ISI · Apr 2022 – Jan 2023

- Organized seminars featuring leading Natural Language Processing researchers.

Conference Leadership | SoCalNLP19 Student Co-Chair

ISI · Jan 2019 – Sep 2019

- Co-organized the 2nd Symposium on NLP in Southern California and managed the peer-review system.

Community Leadership | Leadership Development Institute

Al Akhawayn University · Aug 2014 – Jun 2015

- Collaborated on social empowerment initiatives in rural Moroccan communities.

Teaching & Mentorship | Teaching Assistant

≡ USC / EPFL

- Advanced NLP, Data Structures & Object-Oriented Design, and Software Engineering.

HONORS

- Honors BSc in Computer Science
- Second-highest CS GPA in graduating class
- University merit scholarship
- Bachelor: President's List and Dean's List.

LANGUAGES

Moroccan Darija ●●●●●
Standard Arabic ●●●●●
English ●●●●●
French ●●●●●
Spanish ●●○○○