Sagi Shaier

Boulder, CO | sagishaier@gmail.com | Linkedin.com/in/Shaier | Publications | GitHub.com/Shaier | Website

SUMMARY

I'm an AI researcher passionate about building language models that are efficient, trustworthy, and practical. My work spans both theoretical and applied aspects of question answering, sparsity, transfer learning, continual learning, and factuality, driven by the goal of creating real-world impact.

PROFESSIONAL EXPERIENCE

Cohere · Intern of Technical Staff May 2025-present

- Agents, Embeddings, Search

University of Colorado Boulder · Doctoral Researcher Aug 2020-May 2025

- **Minimized inference requirements by over 90% in "foundational" models** using biologically-inspired algorithms, accelerating progress towards artificial general intelligence (AGI) and highly-scalable generative AI.
- Developed a framework for **building more accurate and trustworthy language models** that prioritize factuality and mitigate hallucinations by grounding their generation with citations.
- **Increased accuracy by up to 42%** in large language models (LLMs) knowledge assessment through the development of a novel evaluation method, which also **reduces language generation redundancy by up to 40%**.
- Managed 10+ groups of graduate students on a variety of natural language processing projects, such as information retrieval, multihop question answering (QA), knowledge graphs, multilingual AI, and dialogue systems.
- **Improved accuracy by up to 36% in retrieval augmented generation (RAG)** models using attention-based strategies.
- **Awarded the prestigious Social Impact Award** out of 363 international candidates for identifying fairness concerns in biomedical QA systems, paving the way for more reliable and fair AI in high-stakes healthcare applications.

National Institute of Mental Health (NIMH) · ML Researcher (Volunteer) May 2023-May 2025

- **Designed a biologically-inspired mixture-of-experts algorithm** to induce sparsity and modularity in any neural system, enhancing model efficiency, performance, and transfer learning capabilities.
- Collaborated with several researchers teams in **developing multimodal NLP applications**, such as knowledge representation and QA systems in the biomedical domain using both structured and unstructured data.
- Produced intuitive visualizations that **simplify complex algorithms**, making them accessible to a broader audience.
- Developed **biologically-inspired continual learning algorithms** for computer vision and language models.

Oracle · Research Intern Jan 2024-April 2024

- Designed and implemented scalable multi-GPU machine learning systems for large scale training and inference, which **supports 70B parameter models** in both supervised and unsupervised settings.
- **Developed 5 novel datasets** to assess language models' ability to answer complex, ambiguous questions with citations, spanning multiple domains and requiring challenging multihop reasoning.
- **Increased LLMs' question answering accuracy by 19.4%** and **citation generation accuracy by 86.7%** through innovative prompt engineering and fine-tuning techniques, yielding more accurate, trustworthy, and reliable AI.

Pacific Northwest National Laboratory (PNNL) · National Security Research Intern May 2021-October 2021

- **Revealed novel patterns in high-dimensional text data and word embeddings** using various techniques intersecting topology and natural language processing, uncovering intrinsic logic and advancing natural language understanding.
- **Improved factual knowledge representation** with non-Cartesian normalization methods.

Quantum Metric · Data Scientist (Research Team) Dec 2019-May 2020

- **Optimized query performance** and analytics on massive datasets by integrating Google's BigQuery, AutoML, and Google cloud computing, and enabling efficient processing of massive datasets.
- **Maximized business outcomes** through predictive behavioral analytics, leveraging deep learning techniques like XGBoost and K-Means to drive revenue growth and customer engagement in diverse industries.
- **Led research initiatives** that applied advanced ML methods to detect real-time app crashes and service disruptions, **saving millions annually** by minimizing downtime and improving customer satisfaction.

Welocalize · Machine Learning Intern May 2019-Aug 2019

- **Grew project management capabilities** by developing predictive regression models using advanced data analysis techniques, complex SQL queries, Pandas, Numpy, and data visualization with Plotly.
- **Delivered actionable insights** through data analysis, utilizing clustering, Seaborn correlations, and dimensionality reduction to identify organizational bottlenecks, and optimizing the time tracking system for improved performance.

Hack Oregon · Data Scientist (Volunteer) Feb 2019-Sep 2019

- Created a spatial distribution map of potential casualties of a Cascadia Earthquake to **inform medical response strategies**.

EDUCATION

University of Colorado Boulder · PhD Computer Science
Dissertation: "Factual Knowledge-enhanced Question Answering in Dynamic Environments".

University of Colorado Boulder · MS Computer Science

Kennesaw State University · BS Computational and Applied Mathematics,
Concentration in Epidemiology, Minor in Statistics, Pre-Med.

PUBLICATIONS

[1] **S. Shaier**, G. Baker, C. Sridhar, K. von der Wense, L. Hunter, and M. Jones. MALAMUTE: A Multilingual, Highly-granular, Template-free, Education-based Probing Dataset (**ACL Findings**) 2025.

[2] **S. Shaier**, F. Pereira, K. von der Wense, L. Hunter, and M. Jones. More Experts Than Galaxies: Conditionally-overlapping Experts With Biologically-inspired Fixed Routing (**ICLR**) 2025.

[3] **S. Shaier**, A. Kobren, and P. Ogren. Adaptive Question Answering: Enhancing Language Model Proficiency for Addressing Knowledge Conflicts with Source Citations. Empirical Methods in Natural Language Processing (**EMNLP**) 2024.

[4] **S. Shaier**, L. Hunter, and K. von der Wense. It Is Not About What You Say, It Is About How You Say It: A Surprisingly Simple Approach For Improving Reading Comprehension. Association for Computational Linguistics (**ACL Findings**) 2024.

[5] **S. Shaier**, L. Hunter, and K. von der Wense. Desiderata For The Context Use Of Question Answering Systems. European Chapter of the Association for Computational Linguistics (**EACL**) 2024.

[6] **S. Shaier**, K. Bennett, L. Hunter, and K. von der Wense. Comparing Template-based And Template-free Language Model Probing. European Chapter of the Association for Computational Linguistics (**EACL**) 2024.

[7] **S. Shaier**, L. Hunter, and K. von der Wense. Who Are All The Stochastic Parrots Imitating? They Should Tell Us!. Asia-Pacific Chapter of the Association for Computational Linguistics (**AACL**) 2023.

[8] **S. Shaier**, K. Bennett, L. Hunter, and K. von der Wense. Emerging Challenges In Personalized Medicine: Assessing Demographic Effects On Biomedical Question Answering Systems. Asia-Pacific Chapter of the Association for Computational Linguistics (**AACL**) 2023. **Won Social Impact Award.**

[9] **S. Shaier**, M. Raissi, and P. Seshaiyer. Data-driven approaches for predicting spread of infectious diseases through DINNs: Disease Informed Neural Networks (**Letters in Biomathematics**) 2022.

HONORS & AWARDS

University of Colorado Boulder · Best Research Poster Award	($500)	Feb 2025
Bell Family Foundation · Outstanding Research Award	($1,000)	Nov 2024
University of Colorado Boulder · Outstanding Research Paper Award	($500)	May 2024
University of Colorado Boulder · Outstanding Service Award	($500)	May 2024
University of Colorado Boulder · Publication Recognition Award (x5)	($4,000)	Aug 2023 - 2024
AACL · Social Impact Award	($0)	Aug 2023
Nelson A. Prager Family Fund · James H Martin Graduate Award	($2,000)	Aug 2023
University of Colorado Boulder · Outstanding Student Award	($500)	Aug 2023

SKILLS

Programming Languages: Python, Bash, SQL, Java, MATLAB
ML Frameworks: PyTorch, TensorFlow, HuggingFace, OpenCV, Keras, NumPy, Pandas, sklearn, scikit-learn, Seaborn, AWS
Tools: Git, Docker, TensorBoard, WandB, OpenSearch, MCP, Linux/Unix, LaTeX, OpenAI
Project Management: strategic planning, budgeting, goal posting, delegation and supervision
Communication: scientific and analytical writing, public speaking and presenting, teaching and training