

DANIAL AMIN

Doctoral Student | Doctoral Researcher

+358 46 801 4291 | writetodanialamin@gmail.com | [linkedin](#) | [portfolio](#) | [github](#) | [scholar](#) | [ORCID](#) | [team](#) | Espoo, Finland

PROFILE SUMMARY

As a PhD candidate, my research combines expertise in data science, machine learning, and large language models to address stereotype leakage and bias in AI-generated user representations, particularly for marginalized communities. I develop demographically accurate and culturally grounded personas using retrieval-augmented generation (RAG) frameworks and open-source LLMs. My work emphasizes inclusive data collection, rigorous evaluation methodologies, and ethical frameworks, ensuring that AI systems authentically represent underrepresented groups and align with the principles of responsible AI.

SKILLS AND COMPETENCIES

Research Interests: AI-Generated Personas for Social Good, Bias Detection and Mitigation in AI Systems, Ethical AI for Marginalized Communities, Data-Driven Persona Creation, Generative AI and Deepfake Technologies, Cross-Cultural AI Applications, Algorithmic Fairness in Persona Development, AI Ethics and Policy Making, Inclusive Design for Global South, Quantitative and Qualitative Persona Evaluation Methods

AI & Data Science: Predictive Machine Learning, Reinforcement Learning, Deep Learning, Deep Neural Networks, RNNs & GANs, Bayesian Optimization, Generative AI, Large Language Models, Explainable AI, Retrieval Augmented Generation, Exploratory Data Analytics, Predictive and Prescriptive Analytics, Time Series Analysis & Forecasting

Programming Languages & Libraries: Python, R, Pytorch, PySpark, TensorFlow, Keras, Pandas, SciPy, Scikit-learn, Matplotlib, Seaborn, NLTK, MATLAB, LabVIEW, LangChain, LangSmith, Docker, SpaCy, Dragonfly, SQL

Software & Developer Tools: Tableau, Power BI, Looker, Minitab, SPSS, Visual Studio, PyCharm, Anaconda, JupyterLab, Docker

Cloud Architecture: Microsoft Azure, Google Cloud Platform, Amazon Web Services

Languages: English (C1), French (A1), Urdu (C2), Punjabi (C1)

EDUCATION

University of Vaasa, Finland

Doctor of Science in Business Administration and Economics

Rennes School of Business

Masters in Business Administration

Massachusetts Institute of Technology

Micro Masters in Statistics and Data Science

University of Wisconsin, Madison

Micro Masters in Business: Data, Insights, and Analytics

University of California San Diego

Micro Masters in Data Science

Rochester Institute of Technology

Design Thinking

Institute of Space Technology

B.S. in Aerospace Engineering

Jan, 2025 – Date

On-Campus, Finland

Sep. 2023 – Sep. 2024

On-Campus, France

Aug. 2019 – Nov. 2023

Distance Learning, USA

Aug. 2022 – May 2023

Distance Learning, USA

Aug. 2021 – Mar. 2023

Distance Learning, USA

Jan. 2022 – Nov. 2023

Distance Learning, USA

Sep. 2013 – Aug. 2017

On-Campus, Pakistan

EXPERIENCE

Research Experience

AI Research Scientist

Samsung Design Innovation Center

Nov. 2024 - Present

On site, France

- Developing the internal design companion for product and computational designers, combining the Generative and Predictive AI solutions to support data-driven design optimization.
- Spearheading development of multi-stages unstructured DB to act as ground truth for the LLM to enhance the data-driven design optimization process at SDIC.

AI & Data Science Lead

i13 Ventures Studio

Sep. 2023 - Nov 2024

Remote, Ireland

- Developed an LLM-based venture innovator, improving product performance by grounding and contextualizing the model through multi-query, re-ranking RAG algorithms.
- Created AI-based financial advisor, optimizing savings and investment plans by 25% through explainable AI advice using observer /observee agentic framework.
- Designed short, medium, and long-term strategy for business incubation using generative AI, enhancing incubation cycle impact by 30% with a focus on explainability.

Lead Data Scientist

Automat Solutions Inc

Feb. 2022 - Feb. 2023

Remote, USA

- Built machine learning pipelines for battery performance prediction, achieving 95% accuracy by accurately estimating the algorithmic uncertainties.
- Optimized battery material performance, increasing throughput by 40%, by using a combination of optimization algorithms.

- Automated ML pipelines, classifying chemical compositions with 90% precision, focusing on explainability and fairness.
- Lead ML Engineer** Feb. 2021 - Feb. 2022
Omdena Inc Remote, USA
- Improved image classification model accuracy by 12%, leading a team of 5 engineers to tackle recall minimization by incorporating HIL feedback.
 - Developed an object recognition pipeline, enhancing detection accuracy by 15%, with a focus on reducing bias in image datasets.

- Assistant Manager** Aug. 2017 - Nov. 2022
Government of Pakistan On-site, Pakistan
- Implemented a fuel consumption estimator, increasing accuracy by 72%, by developing surrogate model.
 - Developed a SMART Maintenance System, optimizing schedules for 52 aircraft components through anomaly detection algorithm in time series data.
 - Introduced AI-enabled early warning systems, reducing flight risks by 27% through risk assessment models.

Management Experience

- AI & Data Science Lead** Sep. 2023 - Nov. 2024
i13 Ventures Studio Remote, Ireland
- Mentored and nurtured startups, increasing implementation of AI solutions by 50%, focusing on ethical AI and bias mitigation in emerging markets of South Africa.
 - Envisioned AI and ML strategy, leading to a 15% improvement in organizational efficiency.

- Head of Data Analytics — Lead Data Scientist** Feb. 2023 - Sep. 2023
Abacus Global On-site, Pakistan
- Improved data processing speed by 40% and predictive analytics model accuracy by 25% through the implementation of a big data analytics framework.
 - Led a cross-functional team to drive data-centric strategies, addressing biases and ensuring fair data practices.
 - Developed solutions that improved customer satisfaction by 30% by addressing biases in data analysis.

Teaching Experience

- Teaching Assistant** Aug. 2016 - Aug. 2017
Institute of Space Technology On-site, Pakistan
- Assisted in Aero-Vehicle Performance and Design courses by providing hands-on tutorials and problem-solving exercises.
 - Supported course material development, enhancing the learning experience for over 100 students by focusing on hands-on and applied content.

PUBLICATIONS

Working Papers

- How Is Generative AI Used for Persona Development?: A Systematic Review of 52 Research Articles** 2025
Amin, D., Salminen, J., Ahmed, F., Tervola, M.H. S., Sethi, S. & Jansen, B., *ACM Transactions on Computer-Human Interactions* [Preprint](#)
[Under Review]
- Toward using personas for inclusive design: The question of quantifying diversity in persona sets** 2025
Amin, D., Salminen, J., Jansen, B., & Sethi, S., *AAAI/ACM Conference on AI, Ethics, and Society* [Under Review] AIES'25
- Generative AI Personas Considered Harmful? Examining the Risks of Algorithmic User Representations in HCI Design** 2025
Amin, D., Salminen, J., Jansen, B., Shin, J., & Kim, D. H., *International Journal of Human-Computer Studies*. [Under Review] IJHCS
- Large language models (LLMs) in human-computer interaction: Using LLM-generated personas to model everything from minority views to entire ecosystems.** 2024
Amin, D., Salminen, J., Jansen, B., Kaate, I., & Akhtar, W., *In Artificial Intelligence & Large Language Models: A Scientific Perspective* [In press] CRC Press
- How Is AI Used for User Representation?: An Analysis of 83 Persona Prompts** 2025
Salminen, J., Amin, D., & Jansen, B., *ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)* [AICCSA'2025](#)
[Under Review]
- Representing Religious Practices via AI-Generated Personas: A Case Study of Ramadan Behaviors from Four Pre-dominantly Muslim Countries** 2025
Alqadi, L., Jung, S., Amin, D., Alabed, A., Salminen, J. & Jansen, B., *ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)* [AICCSA'2025](#)
[Under Review]
- On the Need for Computational Resources for the Scientific Benchmarking of Data-Driven Personas** 2025
Salminen, J., Amin, D., & J., Jansen, B., *Human-Computer Interaction* [Under Review] HCI
- Quality In, Quality Out: Investigating the Impact of Survey Response Quality on Data-Driven Persona Development** 2025
Salminen, J., Amin, D., & Jansen, B., *Conference of the European Survey Research Association (ESRA)* [In publication] ESRA'25
- The Use of Large Language Models in HCI: A Critical Analysis of Synthetic Users** 2025
Salminen, Amin, D., Jung, S., & J., Jansen, B., *Poster at Augmented Humans Conference* AH'2025

Conference Papers

- **“When AI Writes Personas”: Analyzing Lexical Diversity in LLM-Generated Persona Descriptions** 2025
Sethi, S., Salminen, J., Amin, D., & J., Jansen, B., *Extended Abstracts of CHI’2025* [CHI’2025](#)
- **The Role of National Centers in Sustainable Development of Pakistan** 2021
Naqvi, N., Ahmed, F., Ahmed, U., Amin, D., & Malik, F. ,*International Conference on Aerospace Science and Engineering* [IEEXplore](#)
- **Role of Private-Public Partnerships in Creation of Space Awareness and Outreach in Pakistan** 2021
Naqvi, N., Ahmed, F., Ahmed, U., Amin, D., & Malik, F. ,*International Conference on Aerospace Science and Engineering* [IEEXplore](#)
- **Role of subject matter experts in creating awareness through space technology education and popularization (STEP): World space week celebrations — A case study** 2017
Naqvi, N., Ahmed, F., Ahmed, U., & Amin, D.,*International Conference on Aerospace Science and Engineering* [IEEXplore](#)
- **Unification of traditional and scientific pedagogy in Space Technology Education and Popularization (STEP): Space summer school—A case study** 2017
Naqvi, N., Ahmed, F., Ahmed, U., & Amin, D.,*International Conference on Aerospace Science and Engineering* [IEEXplore](#)

Courses

- **Bayesian Machine Learning for Optimization in Python** 2023
Educative Inc [Educative](#)

PROJECTS

PEP-Persona Ecosystem Playground

- Developing an interactive playground to support interaction of representative personas of different stakeholders among themselves and with human-in-loop
- Developing a framework to isolate user representation from the user information
- **Tech Stack** :Python, LangChain, Pinecone, Azure, LangServe, FastAPI

Bob-The Startup Advisor

- Created a selection and scoring agent to assist innovation hubs around the globe in evaluating and selecting startups in emerging markets holistically using LLMs
- Developed and deployed an enterprise level contextualized and grounded large language model for supporting startups through ideation and incubation stage.
- **Tech Stack** :Python, LangChain, Pinecone, Azure, LangServe, FastAPI

Bayesian Optimization Network for Chemical Composition Optimization

- Implemented a project for generating chemical compositions and optimizing the process using targeted DNNs through BO, Active Learning, and Meta-learning models.
- **Tech Stack** :Python, DragonFly, Bayesian Neural Networks, Genetic/Evolutionary Algorithms

AI-Shyr: Where ML meets poetry

- A pre-LLM based poetry generator for Urdu language using corpus of local poets
- **Tech Stack** :Python, Streamlit, RNNs, LSTM, nGrams,Space,NLTK

COURSES & CERTIFICATIONS

- Professional Certificate in **Artificial Intelligence and Machine Learning** by **European IT Certification Academy** [EITCA](#)
- Program in **Data Science and ML: Making Data-Driven Decisions** by **MIT Institute for Data, Systems, & Society** [MIT IDSS](#)
- Professional Certificate in **Data Analytics** by **Google** [Coursera](#)
- Professional Certificate in **Project Management** by **Google** [Coursera](#)
- Professional Certificate in **Data Science** by **IBM** [Coursera](#)
- Specialization in **Deep Learning** by **Deeplearning.ai** [Coursera](#)
- Specialization in **Natural Language Processing** by **Deeplearning.ai** [Coursera](#)
- Specialization in **Reinforcement Learning** by **University of Alberta** [Coursera](#)
- Complete Career track of **Data Scientist** by **365DataScience** [365DataScience](#)
- Expert track in **Data Science Fundamentals on Microsoft Azure** by **CloudSwyft and Microsoft** [FutureLearn](#)
- Nano degree in **AI Programming with Python** by **Udacity** [Udacity](#)
- Developer Certification in **Data Analysis with Python** by **FreeCodeCamp** [FreeCodeCamp](#)

HONORS & AWARDS

International

- **Climate Change Scholar** for **French Excellence Scholarship** by **Campus France** 2023-24
- **Forward Fellow** for **McKinsey Forward Program** by **McKinsey** 2022
- **Next Generation Tech Booster Scholarship** for **Data Science Track** by **Udacity** 2022
- **AWS Scholarship** for **ML Engineering Program** by **Udacity** 2022
- **Honors for Applied Data Science Lab** by **World Quant University** 2022
- **Future Learn Award** for **Thriving in Digital Workplace** by **University of Leeds** 2021
- **Future Learn Award** for **Solving Problem with Technology** by **University of Leeds** 2021
- **Bronze Medal** for **Poster Competition** by **Asian Science Camp** 2015

National

- **National Champion** for **Office Specialization** by **Microsoft** 2016
- **Salutatorian** for **Class of 2017** by **Institute of Space Technology** 2017
- **Cum Laude** for **Class of 2017** by **Institute of Space Technology** 2013-17
- **Outstanding Talent Award** for **Class of 2017** by **Institute of Space Technology** 2017
- **International Achievement Award** for **Class of 2017** by **Institute of Space Technology** 2016-2017

PARTICIPATION & VOLUNTEERING

- Program Committee Member, **WebScience Conference** 2025
- Participant, **Hi Paris Summer School for AI & Data for Science** , Business and Society 2024
- Facilitator, **Space Summer School at National Center of GIS & Space Applications** 2017-23
- Resource Person, **World Space Week** at Institute of Space Technology 2015-23
- Keynote Speaker, 5th - 7th **International Conference on Aerospace Science and Engineering** 2017-21