

Fahad Ahmed Khokhar

(+39) 351-608-4234 | fahadahmed.khokhar@unifi.it | Personal Website | LinkedIn | GitHub | Florence, Italy | Date of Birth: 30-03-1997

SUMMARY

AI Researcher with a PhD in Computer Science from the University of Florence, Italy specializing in Trustworthy AI, Agentic AI, Large Language Models, AI Safety, and autonomous intelligent systems. My research focuses on building reliable AI through uncertainty estimation, runtime safety monitoring, safety wrappers, and fail-controlled machine learning. Experienced in developing multi-agent LLM systems with reasoning, planning, Retrieval-Augmented Generation (RAG), memory, and tool orchestration. Interested in the safety, robustness, evaluation, and certification of autonomous AI agents interacting with real-world software systems.

AWARDS

Best Paper Award

30th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC)

Nov 2025

Seoul, South Korea

RESEARCH INTERESTS

Agentic AI, AI Safety, Trustworthy AI, Large Language Models, Multi-Agent Systems, LLM Evaluation, AI Alignment, Adversarial Machine Learning, Runtime Monitoring, Autonomous Systems

TECHNICAL SKILLS

Programming: Python, C++, SQL, Java, JavaScript

AI & Machine Learning: Large Language Models (LLMs), Agentic AI, AI Safety, Trustworthy AI, Retrieval-Augmented Generation (RAG), Multi-Agent Systems, Transformers, Computer Vision, Federated Learning, Diffusion Models

Safety & Evaluation: Runtime Monitoring, Uncertainty Estimation, Adversarial Machine Learning, LLM Evaluation, Safety Wrappers, Prompt Engineering

Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, Keras, MONAI, OpenCV

Libraries: NumPy, Pandas, Scikit-learn, Matplotlib, Weights & Biases

Tools & Platforms: Docker, Git, Linux, PostgreSQL, MySQL, VS Code, Anaconda

EXPERIENCE

Data Scientist

AIGOT

Sep 2025 – Jan 2026

Pisa, Italy

- Developed Agentic AI systems using LLaMA and Mistral with multi-agent reasoning, memory, RAG, and tool integration.
- Fine-tuned and evaluated open-source LLMs on 50K+ domain-specific samples for Italian-language applications.
- Implemented end-to-end diffusion and generative AI pipelines for text-to-image and content generation workflows.

Research Assistant

COMSATS University Islamabad

Feb 2019 – Jan 2022

Wah Cantt, Pakistan

- Led software development for HEC-sponsored robotic fish project for automated water disease detection
- Implemented computer vision models for real-time bacterial contaminant identification
- Co-authored Elsevier publication on federated learning for image processing

Software Developer

Pakistan Ordnance Factories

Jan 2021 – Jan 2026

Wah Cantt, Pakistan

- Developed and optimized Oracle-based enterprise MIS applications serving 500+ users
- Automated reporting workflows, reducing manual data processing time by 40%

Freelance AI Developer

Upwork

Jun 2019 – Aug 2022

Remote

- Delivered 10+ ML/DL projects including data pipelines, model training, and web scraping tools
- Contributed to academic manuscripts and technical documentation for AI research clients

KEY PROJECTS

Agentic AI Safety for EDA
Trustworthy AI Research

Jan 2026 – Jun 2026
Python, PyTorch, Hugging Face, LLMs, Agentic AI

- Developed a framework for evaluating autonomous LLM agents performing system-level tasks through APIs and multi-step reasoning.
- Studied adversarial threats including prompt injection, agent hijacking, unsafe tool execution, and policy violations in agentic workflows.
- Designed runtime safety monitors based on uncertainty estimation and safety wrappers to improve the reliability of autonomous AI systems.

Food Dish Generation System
Deep Learning

Sep 2025 – Jan 2026
Python, TensorFlow, Hugging Face Diffusers, LLMs

- Built text-to-image generation system using Stable Diffusion for food image synthesis
- Curated custom dataset of 10K+ dish images with LLM-generated captions
- Implemented full pipeline using Hugging Face Diffusers and transformer-based architectures

Fail-Controlled Classifiers
PhD Research – Trustworthy AI

Nov 2022 – Aug 2024
Python, PyTorch, Scikit-Learn

- Developed runtime safety monitoring frameworks enabling trustworthy AI systems to detect and reject uncertain predictions before deployment.
- Proposed fail-controlled classifier architectures evaluated across 8+ image and tabular benchmarks for safety-critical AI applications.
- Published research in Software: Practice and Experience, Scientific Reports, IEEE PRDC, and ACM SAC, including an IEEE Best Paper Award.

Person Re-Identification using Federated Learning
M.S. Thesis – Computer Vision

Aug 2021 – Dec 2021
Python, PyTorch, OpenCV

- Developed privacy-preserving person re-ID system using federated learning across decentralized nodes
- Achieved competitive accuracy while preserving data privacy through distributed training

Human Face Detection and Recognition
Computer Vision

Jul 2017 – Jan 2018
C++, OpenCV, Python, YOLO, TensorFlow

- Built drone-based face detection and recognition system achieving 93% accuracy from aerial angles
- Utilized YOLOv3 with pre-trained weights for real-time detection at 45-degree camera view

EDUCATION

University of Florence
Ph.D. in Computer Science (Trustworthy AI, Safety, Machine Learning)

Florence, Italy
Sep 2022 – Apr 2026

COMSATS University Islamabad
M.S. in Computer Science (Computer Vision, Machine Learning)

Wah Cantt, Pakistan
Feb 2020 – Jan 2022

COMSATS University Islamabad
B.S. in Software Engineering

Wah Cantt, Pakistan
Sep 2014 – Jun 2018

JOURNAL PUBLICATIONS

1. **Khokhar, F. A.**, Zoppi, T., Ceccarelli, A., Montecchi, L., & Bondavalli, A. (2026). *Fail-Controlled Classifiers: A Swiss-Army Knife Toward Trustworthy Systems*. Software: Practice and Experience, 56(3), 239–259. Wiley.
2. **Khokhar, F. A.**, Zoppi, T., Cennini, L., Ceccarelli, A., & Bondavalli, A. (2026). *Safety Monitors for Black-Box Classifiers: Design, Framework and Benchmark*. Scientific Reports.
3. Cheema, M. N., Nazir, A., Harmanci, A., Harmanci, A. S., Cheema, Y., Masood, S., & **Khokhar, F. A.** (2026). *Large Language Models in Radiogenomics: A Comprehensive Survey of Applications from Imaging to Genetics*. The Visual Computer.
4. **Khokhar, F. A.**, Shah, J. H., Saleem, R., & Masood, A. (2025). *Harnessing Deep Learning for Faster Water Quality Assessment: Identifying Bacterial Contaminants in Real Time*. The Visual Computer, 41(2), 1037–1048. Springer.

5. Sayin, B., Zoppi, T., Marchini, N., **Khokhar, F. A.**, & Passerini, A. (2025). *Bringing Machine Learning Classifiers into Critical Cyber-Physical Systems: A Matter of Design*. IEEE Access.
6. Cheema, Y., Cheema, M. N., Nazir, A., **Khokhar, F. A.**, Li, P., & Ahmed, A. (2025). *A Novel Approach for Improving Open Scene Text Translation with Modified GAN*. The Visual Computer, 41(2), 869–881. Springer.
7. Rehman, R., Shah, J. H., **Khokhar, F. A.**, Ahmed, A., Saleem, R., & Najam, S. S. (2025). *Smart Monitoring: Employing Person Re-Identification to Uncover Suspicious Behavior*. Signal, Image and Video Processing, 19(1), 9–20. Springer.
8. Afzal, M., Shah, J. H., ur Rehman, S., **Khokhar, F. A.**, Yasmin, M., & Kadry, S. (2025). *Automated Soccer Event Detection and Highlight Generation for Short and Long Views*. Multimedia Tools and Applications, 84(26), 30971–30991. Springer.
9. Shah, J. H., Afzal, M., Riaz, S., Yasmin, M., Kadry, S., & **Khokhar, F. A.** (2025). *A Comprehensive Dataset of Soccer Event Images for Advancing Automatic Recognition Systems*. Data in Brief, 60, 111518. Elsevier.
10. Fatima, I., Shah, J. H., Saleem, R., Riaz, S., Rafiq, M., & **Khokhar, F. A.** (2024). *Transforming Medical Imaging: A VQA Model for Microscopic Blood Cell Classification*. IEEE Access.
11. Qadeer, N., Shah, J. H., Sharif, M., Dahan, F., **Khokhar, F. A.**, & Ghazal, R. (2024). *Multi-Camera Tracking of Mechanically Thrown Objects for Automated In-Plant Logistics by Cognitive Robots in Industry 4.0*. The Visual Computer, 40(12), 9063–9082. Springer.
12. **Khokhar, F. A.**, Shah, J. H., Khan, M. A., Sharif, M., Tariq, U., & Kadry, S. (2022). *A Review on Federated Learning Towards Image Processing*. Computers and Electrical Engineering, 99, 107818. Elsevier.

CONFERENCE PUBLICATIONS

1. Khokhar, F. A., Zoppi, T., Shah, J. H. et al. (2025). *Orchestrating Fail-Safe, Black-Box Models within Federated Learning Scenarios*. Proceedings of the 30th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2025). (**Best Paper Award**)
2. Zoppi, T., Atif, M., Khokhar, F. A., & Bondavalli, A. (2026). *Back to the Future: Refurbishing Critical Software Engineering to Enable Trustworthy Classification*. ACM Symposium on Applied Computing (SAC 2026).
3. Zoppi, T., Khokhar, F. A., Ceccarelli, A., & Bondavalli, A. (2024). *Position Paper—Bringing Classifiers into Critical Systems: Are We Barking Up the Wrong Tree?*. International Conference on Computer Safety, Reliability, and Security (SAFECOMP 2024). Springer.
4. Zoppi, T., Khokhar, F. A., Montecchi, L., Ceccarelli, A., & Bondavalli, A. (2024). *Fail-Controlled Classifiers: Do They Know When They Don't Know?*. 29th IEEE Pacific Rim International Symposium on Dependable Computing (PRDC 2024). IEEE.
5. Raza, M. R., Ahmed, S., Khokhar, F. A., & Varol, A. (2025). *Exploring the Potential of DeepSeek-R1 Model in Transforming Healthcare Solutions: An Overview*. 2025 13th International Symposium on Digital Forensics and Security (ISDFS). IEEE.

CERTIFICATIONS

- Machine Learning Specialization – Stanford University (Coursera)
- Deep Learning Specialization – DeepLearning.AI (Coursera)

LANGUAGES

English: Fluent (C2) | **Italian:** Basic (A2) | **Urdu:** Native | **Punjabi:** Native