



SAIFULLAH

PhD Candidate · AI Researcher @ DFKI
Kaiserslautern, Germany

(+49) 162-889-3391
saifullah.saifullah@dfki.de
linkedin.com/in/saifullah3396
github.com/saifullah3396
saisai.sh

Summary

PhD researcher specializing in trustworthy AI and generative models for multimodal document understanding. Experienced in building end-to-end AI systems, from large-scale distributed training and evaluation infrastructure to backend services and real-time inference on edge hardware. Current work focuses on designing and deploying LLM-driven agentic workflows.

Skills

AI & Machine Learning

Large Language Models · Document Understanding · Trustworthy AI · Diffusion Models · Computer Vision · Active Learning · Reinforcement Learning

Agentic AI & LLM Systems

Agentic Workflows · Multi-Agent Systems · Tool & API Integration · Retrieval & Context Engineering · Model Context Protocol (MCP) · Pydantic AI / Graph

ML Engineering & Infrastructure

Distributed Training (HPC / SLURM) · Backend & API Design · Data Pipelines & Versioning · Edge Deployment & Model Optimization

Tools & Technologies

PyTorch · Transformers · Diffusers · PEFT / LoRA · Scikit-learn · vLLM · OpenCV · GStreamer · SHAP · Captum · Opacus · Hydra · Weights & Biases · Pandas / NumPy · S3 / LakeFS · DeltaLake · FastAPI · Docker · NVIDIA DeepStream / Jetson · TensorRT / ONNX · Ray · SQLAlchemy · PostgreSQL · MongoDB · RabbitMQ · Qdrant · React / Next.js · Git · Linux · ROS · Claude Code · Codex · GitHub Copilot

Programming Languages

Python ██████████ C/C++ ██████████ JS/TS ██████████
Bash ██████████ LUA ██████████

Experience

- PhD Researcher

German Research Center for Artificial Intelligence (DFKI)

07.2021 — Present
Kaiserslautern, Germany

 - Developed methods to improve the privacy, explainability, and robustness of document AI models for real-world deployment, contributing to 15+ peer-reviewed publications
 - Built and open-sourced the TorchXAI and Atria frameworks to standardize training, evaluation, and explainability across multimodal document AI workflows
 - Developed generative AI methods for controllable synthetic document generation, addressing data scarcity, diversity, and privacy constraints in downstream document understanding applications
- Technical Lead / Research Associate

National Center of Artificial Intelligence (NCAI)

07.2019 — 07.2021
Islamabad, Pakistan

 - Delivered real-time vision solutions built around NVIDIA DeepStream, TensorRT, and Jetsons, enabling pedestrian traffic analytics and automated media trend tracking
 - Delivered an interactive customer support solution using the Pepper humanoid robot for direct user engagement in a banking environment
 - Led the development of a multi-robot navigation system for Kobuki robots, enabling autonomous mapping, obstacle avoidance, and coordinated path planning for simulated warehouse operations
 - Led the development of a UAV simulation environment combining Microsoft AirSim, ROS/PX4, and OpenAI Gym for RL-based autonomous navigation in dynamic environments
- Team Lead Robotics (Part-time)

Ingenio

05.2020 — 09.2020
Islamabad, Pakistan

 - Led a robotics engineering team developing AI-based perception and control systems with custom hardware integration and edge deployment

Selected Projects	AtriaX atriaml.github.io Web-based document AI platform built on Atria for dataset and model management, training, evaluation, and explanation visualization
	TorchXAI atriaml.github.io Explainable AI toolkit with 20+ explainability metrics and 11+ multi-output explainers, enabling fast, scalable evaluation of interpretability in multimodal deep learning models
	DocDjinn arxiv.org General-purpose framework built on vision-language models and diffusion-based handwriting models for controllable generation of full-page synthetic documents, along with their annotations, to train models across diverse downstream document understanding tasks See more on my website: saisai.sh/projects
Selected Publications	Mistaking Periodicity for Manipulation: JPEG-Induced Structural Bias in Document Tampering Detection British Machine Vision Conference, BMVC, 2026 · Co-First Author
	DP-DocLDM: Differentially Private Document Image Generation Using Latent Diffusion Models International Conference on Document Analysis and Recognition, ICDAR, 2025 · 1st Author
	DocVCE: Diffusion-based visual counterfactual explanations for document image classification arXiv, 2025 · 1st Author
	The privacy-explainability trade-off: unraveling the impacts of differential privacy and federated learning on attribution methods Frontiers in Artificial Intelligence, 2024 · 1st Author
	The reality of high performing deep learning models: A case study on document image classification IEEE Access, 2024 · 1st Author
	DocXclassifier: towards a robust and interpretable deep neural network for document image classification International Journal on Document Analysis and Recognition, IJDAR, 2024 · 1st Author
	Coldbin: Cold diffusion for document image binarization International Conference on Document Analysis and Recognition, ICDAR, 2023 · 1st Author
	Are Deep Models Robust against Real Distortions? A Case Study on Document Image Classification International Conference on Pattern Recognition, ICPR, 2022 · 1st Author Full publication list: Google Scholar
Education	Ph.D., Computer Science RPTU Kaiserslautern-Landau 2021 — ongoing Kaiserslautern, Germany
	M.S., Robotics & Intelligent Machine Engineering National University of Sciences and Technology 2016 — 2019 Islamabad, Pakistan Graduated with Honors, Gold Medal
	B.E., Mechanical Engineering National University of Sciences and Technology 2012 — 2016 Islamabad, Pakistan Graduated with Honors
Academic Exchange	Summer School Student Oxford ML School (OxML) 05.2023 Oxford, UK
	Research Exchange Student Sakura Science Exchange Program (SSP) 07.2017 Sendai, Japan
	TeamNUST Member Standard Platform League (SPL), RoboCup 2016 06.2016 Leipzig, Germany
Languages	English (C2) · Urdu (Native) · German (A1)