

RISHABH JAIN

| Email: j_rishabh@outlook.com | Address: Dublin, Ireland |

| Website: <https://rishabhjain16.github.io/> | LinkedIn: <https://www.linkedin.com/in/rishabhjain16> |

| GitHub: <https://www.github.com/rishabhjain16> | HuggingFace: <https://huggingface.co/rishabhjain16> |

DATA SCIENTIST | AI ENGINEER | ML ENGINEER | AI ETHICIST

I am a Data Scientist, AI Engineer, and ML Engineer with a PhD in AI for Speech from the University of Galway, Ireland. My expertise includes Speech-to-Text (STT), Text-to-Speech (TTS), Vision Models, and Large Language Models (LLMs). I have experience in AI ethics, compliance, and governance, particularly in GDPR, risk assessment, and responsible AI. I have collaborated with Data Protection Officers (DPOs) and regulatory teams to ensure privacy, security, and ethical AI deployment. Passionate about trustworthy AI, I strive to build AI technologies that ensure transparency, fairness, and responsible AI solutions.

STRENGTHS AND TECHNICAL EXPERTISE

Soft skills:

Teamwork, Communication,
Leadership, Research work,
Problem Solving, Policy Analysis,
Ethical AI Advocacy, Responsible AI

Languages and ML Libraries:

SQL, R, Python, C, C++, C#, Java, HTML,
CSS, NodeJS, Pandas, NumPy, SciPy,
Librosa, Transformers, Sklearn, Seaborn,
Tensorflow and PyTorch.

Softwares and Cloud Platforms:

Anaconda, Spyder, Jupyter Notebook,
Visual Studio, Eclipse, AWS, GCP,
Docker, Git, Bitbucket

Languages:

English, Hindi, and Spanish.

PROFESSIONAL EXPERIENCE

November 2024 - Present

Trinity College Dublin, Ireland

Research Fellow (Post Doctoral Researcher)

- Collaborating on the groundbreaking Speech Space project, focusing on advancing multimodal audio-video tools.
- Leading the development of multimodal audio-video speech technology utilising Large Language Models (LLMs) to seamlessly integrate multiple modalities, ensuring both accuracy and ethical AI application.
- Conducting pioneering research in visual speech recognition (VSR) technologies, leveraging LLMs to enhance lip-reading capabilities and integrate visual cues with audio for comprehensive speech understanding while adhering to principles of transparency and fairness.

Oxford Wave Research, Oxford, UK

February 2024 - May 2024

Research Scientist (Contract)

- Conduct research on training the models in the domain of Speech-To-Text and LLMs for low resource languages.
- Working with engineers to train and deploy models for production, and optimizing models for real-world use cases such as improving ASR for low resource Albanian speech, which led to improved transcription accuracy by 20%.
- Proficient in finetuning, and deploying advanced LLMs (eg. Llama, Gemma, Gemini, Command-R, Mistral). Leveraged diverse LLM tools and frameworks such as Ollama, Langchain, LM Studio and llama.cpp.
- Creating a pipeline combining Whisper ASR, speaker diarization, speaker segmentation and LLM summarization, enhancing model capabilities for nuanced tasks in NLP and NLU for law and forensics use cases, cutting down manual summarisation time by 40%, thereby boosting productivity for legal and forensic professionals.
- Experienced in LLM Quantization methodologies, implementing efficient model compression techniques to optimize inference speed and resource utilization while maintaining model fidelity.
- Conducted evaluations of LLMs across multiple linguistic domains, leading to precise optimizations that improved model performance, accuracy, and applicability in diverse, real-world tasks critical to legal and forensic analysis.
- Managed and analyzed sensitive forensic speech data, including law enforcement and criminal data, ensuring responsible and ethical AI deployment in forensic investigations.
- Developed models and methodologies to detect and mitigate deep fakes, ensuring the integrity and reliability of AI applications in forensic and legal contexts.

Xperi Galway (now part of Tobii), Ireland

October 2020 - January 2024

Researcher (Consultant)

- Designed and deployed edge AI models (STT and TTS) in collaboration with Xperi, significantly enhancing child-user experience by enabling seamless real-time speech interaction in the DAVID smart-toy.
- Collaborated with cross-functional teams, including engineers, linguists, and product managers from Xperi Ireland and Xperi USA for optimising AI models and deployment in production environment.
- Developed new methodologies for ethical and responsible data collection and annotation of child speech (automatic labelling), addressing challenges like data scarcity and noise, improving the robustness of child speech ASR and TTS models.
- Developed synthetic child speech datasets, resulting in a 5-10% increase in the accuracy of child ASR systems.

- Implemented data anonymisation techniques for child speech datasets, enhancing privacy protections and ensuring compliance with stringent ethical standards and regulations.

October 2020 - January 2024

University of Galway, Ireland

Research Assistant

- Project: DTIF-DAVID, Aim: Development of a multimodal (sound and vision) AI processing platform with low cost and low power consumption to be used for the creation of voice-enabled toys.
- Conduct a comprehensive review of existing research and literature related to child speech technologies (ASR/TTS).
- Developed methodologies for ASR and TTS with child speech, from training to evaluation of models, resulting in synthetic speech datasets that upheld privacy standards and GDPR standards, which also led to the creation of open-source resources and speaker recognition models tailored for various accents in English, spoken by children.
- Developing applications for speech data collection, creating methodologies for data annotations and cleaning.
- Attending conferences, workshops, seminars, and knowledge-sharing activities within research community.
- Documentations of research, preparing reports/presentations and contributing towards novel publications.
- Setting up and managing distributed training pipelines (leveraging multiple GPUs), coordinating data preprocessing and augmentation, monitoring training progress, and evaluating model performance using appropriate metrics.
- Led the adaptation of Whisper models to better recognize child speech, achieving between 20% - 50% improvement in accuracy over different child speech datasets.
- Collaborated with Data Protection Officers (DPOs) to ensure the ethical collection and use of child speech data.
- Developed and implemented policies related to data collection, risk assessment, and compliance with privacy regulations, prioritizing the protection and ethical handling of sensitive child speech data.

EDUCATION

PhD in Artificial Intelligence for Speech

September 2021 - May 2024

University of Galway, Ireland

Research Area: Child speech understanding and generation via neural ASR and TTS models

MSc in Data Analytics

September 2019 - August 2020

University of Galway, Ireland

Grade: 1.1 Honours

Thesis: Semi automatic toolkit for analysis of video, audio and text data based on face recognition in videos.

B.Tech in Computer Science and Engineering with specialization in Bioinformatics

July 2015 - May 2019

VIT University, Tamil Nadu, India

CGPA: 8.73 (equivalent to 1.1 Honours)

PUBLICATION AND CONFERENCES

Publications:

- R. Jain, M. Y. Yiwere, D. Bigioi, P. Corcoran and H. Cucu, "A Text-to-Speech Pipeline, Evaluation Methodology, and Initial Fine-Tuning Results for Child Speech Synthesis," in IEEE Access, vol. 10, pp. 47628-47642, 2022, doi: 10.1109/ACCESS.2022.3170836.
- R. Jain, A. Barcovschi, M. Y. Yiwere, D. Bigioi, P. Corcoran and H. Cucu, "A WAV2VEC2-Based Experimental Study on Self-Supervised Learning Methods to Improve Child Speech Recognition," in IEEE Access, vol. 11, pp. 46938-46948, 2023, doi: 10.1109/ACCESS.2023.3275106.
- Jain, R., Barcovschi, A., Yiwere, M., Corcoran, P., Cucu, H. (2023) Adaptation of Whisper models to child speech recognition. Proc. INTERSPEECH 2023, 5242-5246, doi: 10.21437/Interspeech.2023-935

Conferences:

- Attended IEEE Region 8 SYP congress held in Tunis, Tunisia to represent the University of Galway Student Branch as Chair.
- Worked as a Reviewer for the IEEE Access Journal, IMVIP 23 conference and Sped 23 Conference.
- Organised a special session at the SpeD 23 conference on Child Speech Technologies held in Bucharest, Romania.
- Attended and presented at the Interspeech 2023 in Dublin, Ireland.

CERTIFICATIONS

- AI for Medical Diagnosis by Deeplearning.ai, May 2020
- Introduction to Cyber Security Specialization by NYU, May 2019
- Machine Learning with TensorFlow on GCP Specialization | June 2018
- Deep Learning Specialization by Deeplearning.ai , April 2018

REFERENCES

- Dr. Anil Alexander, CEO, Oxford Wave Research, Oxford, UK, anil@oxfordwaveresearch.com
- Dr. Peter Corcoran, IEEE fellow, Professor at the University of Galway, Lead PI in DAVID, HELIUS and d-real projects, peter.corcoran@universityofgalway.ie
- Dr. Gabriel Costache, Senior Director of Engineering at Tobii (previously part of Xperi Galway), gabriel.costache@tobii.com