

Morrie Doulaty

Work Experience

2026/5–Now **Staff Research Engineer**, *Google DeepMind*, London, UK
- Worked on Speech LLMs and fully duplex voice agents
- Worked on agentic audio systems, modeling, evaluation and orchestration

2021/5–2026/5 **Tech Lead Staff AI/ML Software Engineer**, *Meta (Facebook)*, London, UK
- Led dialogue quality metrics team for natural dialogue systems with audio-in/audio-out interface for Llama family of LLMs
- Led training infra for Ego-centric multi-modal systems for conversational understanding

2020/3–2021/5 **Senior Research Data Scientist**, *Meta (Facebook)*, London, UK
- Led projects focused on designing and implementing large-scale information retrieval systems that utilise neural embeddings for semantic content matching and copy detection
- Implemented AI-based solutions to combat misinformation on Meta's platforms, including Facebook, Instagram, and WhatsApp

2018/6–2020/3 **Senior Applied Scientist**, *Microsoft*, Munich, Germany
- Developed state-of-the-art solutions for speech and language problems within the Cognitive Services domain with deep learning
- Designed and trained ML models for language understanding and speech recognition

2017/12–2018/6 **Research Software Development Engineer**, *Microsoft Research*, Munich, Germany
Collaborated with Microsoft Office teams to deploy advanced conversational AI models

2017/02–2017/12 **Data Scientist**, *QuantumBlack, AI by McKinsey & Company*, London, UK
Developed time series models utilising deep neural networks to perform predictive maintenance tasks

2017–2020 **Visiting Researcher**, *Speech and Hearing Group, University of Sheffield*, Sheffield, UK

2015/07–2015/10 **Research Intern**, *Google*, New York, US
- Worked on multi-style training of deep neural networks for automatic speech recognition
- Developed new techniques for automatic optimisation of data set perturbation distributions (see [Doulaty, Rose and Siohan, 2016] for further details) for training of acoustic models

2013/02–2017/01 **Natural Speech Technology Researcher**, *University of Sheffield*, Sheffield, UK
- Primarily researched multi-domain speech recognition systems with AI and ML techniques
- Shared responsibility for managing a scientific computing cluster and infrastructure and training large-scale AI models for speech recognition

Education

2013–2017 **Ph.D. of Computer Science**, *Speech Technology and Machine Learning*, Speech and Hearing Research Group, Department of Computer Science, University of Sheffield, UK
Thesis Title: Methods for Addressing Data Diversity in Automatic Speech Recognition
Supervisor: Prof. Thomas Hain, Advisor: Prof. Roger K. Moore

2009–2012 **M.Sc. of Computer Science**, *Intelligent Systems*, Department of Computer Science, University of Tabriz, Iran
Thesis Title: Document Classification Using Kernel Approaches

Cambridge – UK
+44 7447 055 999 • *m@doulaty.com*

2005–2009 **B.Sc. of Computer Science (Hons)**, Department of Computer Science, University of Tabriz, Iran
Final Project's Title: Automatic Number Plate Recognition for Arabic Script Letters

Interests

- Artificial Intelligence
- Machine Learning
- Speech and Language Technology
- Data Science

Professional Skills

Solid background in programming, algorithms, and AI in the context of research and industry:

- Proficient in Python, C, C++, C#, and Go programming languages
- Expertise in parallel, grid, and cloud computing, utilising platforms such as Spark, SGE, Slurm, Microsoft Azure, Amazon AWS, and Google Cloud Platform
- Expertise in scientific computing, particularly in Python and NumPy
- Familiar with deep learning frameworks including PyTorch, TensorFlow, and CNTK
- Familiarity with container-based technologies such as Docker and Kubernetes
- Experience in continuous integration and continuous deployment methodologies
- Familiar with databases (SQL & NoSQL) and web technologies (HTML, Django, Flask, ASP .NET)

Selected Publications

- Seide F., **Doulaty, M.**, et al. (2024) Speech ReaLLM – Real-time Streaming Speech Recognition with Multimodal LLMs by Teaching the Flow of Time. *Interspeech 2024*.
- Liu X., . . . , **Doulaty, M.**, et al. (2023). SynthVSR: Scaling up Visual Speech Recognition with Synthetic Supervision. *CVPR 2023*.
- Grauman K., . . . , **Doulaty, M.**, et al. (2022). Ego4D: Around the World in 3,000 Hours of Egocentric Video. In *Proc. of CVPR*.
- **Doulaty, M.**, Hain, T. (2019). Latent Dirichlet Allocation Based Acoustic Data Selection for Automatic Speech Recognition. In *Proc. of Interspeech*.
- Deena, S., Hassan M., **Doulaty, M.**, Saz, O., Hain, T. (2018). Recurrent Neural Network Language Model Adaptation for Multi-Genre Broadcast Speech Recognition and Alignment. *IEEE Transactions on Audio, Speech and Language Processing*.
- Saz, O., Deena, S., **Doulaty, M.**, et al. (2018). Lightly Supervised Alignment of Subtitles on Multi-Genre Broadcasts. *Multimedia Tools and Applications*.
- **Doulaty, M.**, (2017). Methods for Addressing Data Diversity in Automatic Speech Recognition. *Ph.D. thesis. The University of Sheffield*.
- **Doulaty, M.**, Rose, R., Siohan, O. (2016). Automatic Optimisation of Data Perturbation Distributions for Multi-Style Training in Speech Recognition. In *Proc. of SLT*.
- **Doulaty, M.**, Saz, O., Ng, R., Hain, T. (2016). Automatic Genre and Show Identification of Broadcast Media. In *Proc. of Interspeech*.
- Deena, S., Hasan, M., **Doulaty, M.**, Saz, O., Hain, T. (2016). Combining Feature and Model-Based Adaptation of RNNLMs for Multi-Genre Broadcast Speech Recognition. In *Proc. of Interspeech*.

- Hain, T., . . . , **Doulaty, M.**, et al. (2016). webASR 2 - Improved Cloud Based Speech Technology. In *Proc. of Interspeech*.
- **Doulaty, M.**, Saz, O., Ng, R., Hain, T. (2015). Latent Dirichlet Allocation Based Organisation of Broadcast Media Archives for Deep Neural Network Adaptation. In *Proc. of ASRU*.
- **Doulaty, M.**, Saz, O., Hain, T. (2015). Unsupervised Domain Discovery using Latent Dirichlet Allocation for Acoustic Modelling in Speech Recognition. In *Proc. of Interspeech*.
- **Doulaty, M.**, Saz, O., Hain, T. (2015). Data-selective Transfer Learning for Multi-Domain Speech Recognition. In *Proc. of Interspeech*.
- Saz, O., **Doulaty, M.**, Deena, S., Milner, R., Ng, R., Hasan, M., Liu, Y., Hain, T., (2015). The 2015 Sheffield System for Transcription of Multi–Genre Broadcast Media. In *Proc. of ASRU*.
- Milner, R., Saz O., Deena S., **Doulaty, M.**, Ng, R., Hain, T., (2015). The 2015 Sheffield System for Longitudinal Diarisation of Broadcast Media. In *Proc. of ASRU*.
- Loweimi, E., **Doulaty, M.**, Barker, J., Hain, T., (2015). Long-term Statistical Feature Extraction from Speech Signal and its Application in Emotion Recognition. In *Proc. of SLSP*.
- Ng, R., . . . , **Doulaty, M.**, Lee, T., Hain, T., (2015). Sheffield LRE 2015 System Description. In *Proc. of Odyssey: The Speaker and Language Recognition Workshop*.

Awards, Honors and Distinctions

2024 **Recognised as an Exceptional Talent under UK's Global Talent programme**, *endorsed by TechNation*

2022 **Egocentric Vision (EgoVis) 2022/2023 Distinguished Paper Award**, *Awarded for Ego4D: Around the World in 3,000 Hours of Egocentric Video (CVPR 2022) paper*

2013 **Granted EPSRC-funded 3.5 year PhD scholarship**, *Natural Speech Technology Project*, (EP/I031022/1)

Languages

- English: Full professional proficiency
- Azerbaijani Turkish: Native
- Persian (Farsi): Bilingual proficiency
- German: Elementary proficiency

Community Service / Volunteer Work

- Served as a reviewer and area-chair for various conferences and workshops in the field of ML, speech technology and computer vision, including the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Interspeech, Computer Vision and Pattern Recognition (CVPR), IEEE Automatic Speech Recognition and Understanding (ASRU), IEEE Spoken Language Technology (SLT)
- Scholarpedia: Participated as assistant editor
- Manning Publications: Proposal and book reviewer
- Experts-Exchange: Contributed to solving programming problems, have been recognised as a qualified expert