



The  
University  
Of  
Sheffield.

Department  
Of  
Computer  
Science.

|                          |  |
|--------------------------|--|
| <b>Title of Project:</b> | PhD Position in Deep Cascaded Representation Learning for Speech Modelling |
|--------------------------|--|

|                    |                       |
|--------------------|-----------------------|
| <b>Supervisor:</b> | Professor Thomas Hain |
|--------------------|-----------------------|

|                                   |                 |
|-----------------------------------|-----------------|
| <b>Deadline for Applications:</b> | 13th April 2023 |
|-----------------------------------|-----------------|

The LivePerson Centre for Speech and Language offers a 3 year fully funded PhD studentship covering standard maintenance, fees and travel support, to work on cascaded deep learning structures to model speech. The Centre is connected with the Speech and Hearing (SpandH) and the Natural Language Processing (NLP) research groups in the Department of Computer Science at the University of Sheffield.

Auto-encoding is a powerful concept that allows us to compress signals and find essential representations. The concept was expanded to include context, which is usually referred to as self-supervised learning. On very large amounts of speech data this has led to very successful methods and models for representing speech data, for a wide range of downstream processes. Examples of such models are Wave2Vec or WaveLM. Use of their representations often requires fine-tuning to a specific task, with small amounts of data. When encoding speech, it is desirable to represent a range of attributes at different temporal specificity. Such attributes often reflect a hierarchy of information.

The aim in this PhD project is to explore the use of knowledge about natural hierarchies in speech in cascaded auto- and contextual encoder/decoder models. The objective is to describe a structured way to understand such hierarchies. The successful candidate is expected to propose methods to combine different kinds of supervision (auto, context, label) and build hierarchies of embeddings extractions. These propositions may have to be seen in the context of data availability and complexity. All proposals are to be implemented and tested on speech data. Experiments should be conducted on a range of speech data sets with different speech types and data set size.

The student will join a world-leading team of researchers in speech and language technology. The LivePerson Centre for Speech and Language Technology was established in 2017 with the aim to conduct research into novel methods for speech recognition and general speech processing, including end to end modelling, direct waveform modelling and new approaches to modelling of acoustics and language. It has recently extended its research remit to spoken and written dialogue. The Centre hosts several Research Associates, PhD researchers, graduate and undergraduate project students, Researchers and Engineers from LivePerson, and academic visitors. Being fully connected with SpandH brings collaboration, and access to a wide range of academic research and opportunities for collaboration inside and outside of the University. The Centre has access to extensive dedicated computing resources (GPU, large storage) and local storage of over 60TB of raw speech data.

The successful applicant will work under the supervision of Prof. Hain who is the Director of the LivePerson Centre and also Head of the SpandH research group. SpandH was and is involved in a large number of national and international projects funded by national bodies and EU sources as well as industry. Prof. Hain also leads the UKRI Centre for Doctoral Training In Speech and Language Technologies and their Applications (<https://slt-cdt.ac.uk/>) - a collaboration between the NLP research group and SpandH. Jointly, NLP and SpandH host more than 110 active researchers in these fields.

This project will start as soon as possible.

*If English is not your first language, you must have an IELTS score of 6.5 overall, with no less than 6.0 in each component.*

**How to Apply:**

All applications must be made directly to the University of Sheffield using the Postgraduate Online Application Form.

Information on what documents are required and a link to the application form can be found here - <https://www.sheffield.ac.uk/postgraduate/phd/apply/applying>

On your application, please name Prof. Thomas Hain as your proposed supervisor and include the title of the studentship you wish to apply for.

Your research proposal should:

- Be no longer than 4 A4 pages, including references
- Outline your reasons for applying for this studentship
- Explain how you would approach the research, including details of your skills and experience in the topic area

If you have any queries, please contact [phd-compsci@sheffield.ac.uk](mailto:phd-compsci@sheffield.ac.uk)

**Funding Details:**

This position is fully funded by LivePerson, covering all tuition fees and a stipend at the standard UKRI rate.