

# ADAPT Undergraduate Internship Programme 2018

## PROJECT DESCRIPTION

<b>Institution/Team:</b>	ADELE – Social Dialog Agent (Trinity College)
<b>Project Title:</b>	Identification and Annotation of Topic Shift and Topic Shading in Switchboard Dataset to train a neural network speech dialog agent
<b>Suitable for students who are studying in the following areas:</b>	<ul style="list-style-type: none"> <li>• Linguists</li> <li>• Computational Linguistics Students</li> <li>• Computer Science Students (With Interest in Linguistics or Dialogue Agents)</li> <li>• People with a general interest in Linguistics, Dialogue Agents, Machine Learning</li> </ul>
<b>Skills needed:</b>	<ul style="list-style-type: none"> <li>• Process orientated individuals who can learn to reliably recognise topic shift and topic shading in an existing corpus of annotated 10 minute conversations – full training will be given</li> <li>• Ability to follow existing Annotation Standards</li> <li>• Microsoft Excel (for the annotation)</li> </ul>
<b>Project Description:</b>	<p>ADELE is a Personalized Intelligent Companion designed to engage with users through spoken dialog to help them explore topics of interest. The system will maintain a user model of information consumption habits and preferences in order to (1) personalize the user's experience for ongoing interactions, and (2) build the user-machine relationship to model that of a friendly companion.</p> <p>The next phase of the research project aims to tackle three questions:</p> <ul style="list-style-type: none"> <li>• Identifying the right time for ADELE to interrupt a user?</li> <li>• Identifying what form each interruption is rendered?</li> <li>• How do we interrupt the user?</li> </ul> <p>To do this we need a to train our neural network on specific examples of topic shift [1] and topic shading [2] so that we can begin to tackle the above three questions.</p> <p>The aim of this project is to identify and annotate topic shift in an existing corpus of spoken dialog conversations called Switchboard [3].</p> <p>Switchboard is a collection of circa 2400 telephone conversations between two people which is held by the Linguistic Data Consortium. The dataset has already been transcribed and annotated at the dialogue level.</p> <p>The aim of this project is to add another layer of annotation to this at the topic level. This will involve identifying when the topic of conversation shifts or shades and using the provided annotation scheme to mark the up appropriately.</p> <p>[1] Topic Shift: The introduction of a new topic or 'aboutness' of a conversation which is unrelated to the previous topic.</p> <p>[2] Topic Shading: The introduction of a new topic by first establishing its relevance or connection to the topic that is already being discussed. The new topic is usually related to the old topic in a least one way.</p> <p>[3] <a href="https://catalog ldc.upenn.edu/ldc97s62">https://catalog ldc.upenn.edu/ldc97s62</a></p> <p>[4] <a href="http://www.aclweb.org/anthology/W17-7405">http://www.aclweb.org/anthology/W17-7405</a></p>
<b>The Role of the student &amp; benefits gained from participation in this project:<sup>1</sup></b>	<p>The Research Intern will:</p> <ul style="list-style-type: none"> <li>• Gain first-hand experience of the work involved in annotating datasets</li> <li>• Gain experience using XML and PRAAT, an program for the analysis and reconstruction of speech signals.</li> <li>• Gain experience of training and testing a model which may aid identifying topic shift</li> <li>• Test and contribute to the development of the annotation schema to be used</li> <li>• Meet with the rest of the team as required to discuss the annotation and resolve any problems</li> </ul>

<sup>1</sup> ***This is an initial description of the role of the student and it is liable to change following discussions between the investigators and the student.***

	<ul style="list-style-type: none"> <li>• Document progress and work process</li> <li>• See the end product being used to train a Neural Network Dialogue Agent</li> <li>• Get an overview of the research being conducted in ADAPT</li> <li>• Contribute a useful resource to the research community</li> <li>• Attend weekly project meeting as and when required</li> <li>• Get a fuller understanding of the complexities of speech communication and the current state of the art and difficulties in emulating it</li> <li>• Gain a top level understanding of how neural networks for spoken dialogue agents work</li> </ul>								
<b>Who will be working with you?</b>	<p>The Research Intern will be working closely with the ADELE team. This includes Professor Vincent Wade, Chair of Computer Science TCD and Director of the ADAPT Research Centre who's research has mainly focused on Personalisation and Adaptive Hypermedia, Professor Ben Cowan of UCD School of Information &amp; Communication Studies whose research focus is HCI, Christian Saam, a Research Fellow working in speech recognition and processing, Leigh Clarke, a Research Fellow working in HCI, Emer Gilmartin a Research Assistant / PhD student working on Spoken Dialogue and Dialogue Agents, and Brendan Spillane, a PhD student working in the area of Personalisation and the impact of News website design on Credibility</p> <p>The student will gain first-hand knowledge of each member of the team and will attend weekly meeting as and when needed.</p>								
<b>Short description of the group:</b>	<p>2x Professors 2 x Research Fellows 1 x Research Assistant / PhD Student 1 x PhD Student</p> <p>The ADELE research group is part of the ADAPT Centre, an SFI funded Research Centre spread between TCD, DIT, UCD and DCU <a href="https://www.adaptcentre.ie/">https://www.adaptcentre.ie/</a></p>								
<b>Recommended Reading Material:</b>	<p>Some recent publications from the research group</p> <ul style="list-style-type: none"> <li>• Gilmartin, E., Spillane, B., O'Reilly, M., Saam, C., Su, K., Cowan, B. R., ... Campbell, N. (2017). Annotation of greeting, introduction, and leavetaking in dialogues. In <i>Proceedings of the 13th Joint ISO-ACL Workshop on Interoperable Semantic Annotation (ISA-13)</i>. <a href="http://www.aclweb.org/anthology/W17-7405">www.aclweb.org/anthology/W17-7405</a></li> <li>• Gilmartin, E., Spillane, B., O'Reilly, M., Su, K., Saam, C., Cowan, B. R., ... Wade, V. (2017). Dialog Acts in Greeting and Leavetaking in Social Talk. In <i>Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents</i> (pp. 29–30). New York, NY, USA: ACM. <a href="https://doi.org/10.1145/3139491.3139493">https://doi.org/10.1145/3139491.3139493</a> <a href="https://dl.acm.org/citation.cfm?id=3139493">https://dl.acm.org/citation.cfm?id=3139493</a></li> <li>• Spillane, B., Gilmartin, E., Saam, C., Su, K., Cowan, B. R., Lawless, S., &amp; Wade, V. (2017). Introducing ADELE: A Personalized Intelligent Companion. In <i>Proceedings of the 1st ACM SIGCHI International Workshop on Investigating Social Interactions with Artificial Agents</i> (pp. 43–44). New York, NY, USA: ACM. <a href="https://doi.org/10.1145/3139491.3139492">https://doi.org/10.1145/3139491.3139492</a> <a href="https://dl.acm.org/citation.cfm?id=3139492">https://dl.acm.org/citation.cfm?id=3139492</a></li> </ul>								
<b>Other information:</b>									
<b>For further details on this project please contact:</b>	<table border="1"> <tr> <td><b>Name:</b></td> <td>Brendan Spillane</td> </tr> <tr> <td><b>Phone:</b></td> <td></td> </tr> <tr> <td><b>E-Mail:</b></td> <td><a href="mailto:brendan.spillane@adaptcentre.ie">brendan.spillane@adaptcentre.ie</a></td> </tr> <tr> <td><b>Website:</b></td> <td><a href="http://www.adaptcentre.ie">www.adaptcentre.ie</a></td> </tr> </table>	<b>Name:</b>	Brendan Spillane	<b>Phone:</b>		<b>E-Mail:</b>	<a href="mailto:brendan.spillane@adaptcentre.ie">brendan.spillane@adaptcentre.ie</a>	<b>Website:</b>	<a href="http://www.adaptcentre.ie">www.adaptcentre.ie</a>
<b>Name:</b>	Brendan Spillane								
<b>Phone:</b>									
<b>E-Mail:</b>	<a href="mailto:brendan.spillane@adaptcentre.ie">brendan.spillane@adaptcentre.ie</a>								
<b>Website:</b>	<a href="http://www.adaptcentre.ie">www.adaptcentre.ie</a>								