

# CNGL Undergraduate Students as Researchers Programme 2011 **PROJECT DESCRIPTION**

<b>Institution/Track:</b>	Dublin City University	
<b>Project Title:</b>	Arabic Opinion Mining and Sentiment Analysis.	
<b>Suitable for students who are studying in the following areas:</b>	Computer Science and Computational Linguistics.	
<b>Skills needed:</b>	Computer Science and a good knowledge of Arabic grammar would be ideal.	
<b>Project Description:</b>	<p>Opinion mining aims to identify and extract the sentiment of a speaker or a writer with respect to some topic. Nowadays, Opinion Mining technologies are used in advertisement online systems, question answering systems, recommendation systems, etc. This model can be described as follows:</p> <ol style="list-style-type: none"> <li>1) First we identify sentiments using positive/negative/neutral bag of words: e.g., great, amazing, wonderful, bad, and poor.</li> <li>2) Then we locate the subject and object present in a sentence using parsing technologies and relation extraction</li> <li>3) Finally, we classify texts as expressing a positive or a negative sentiment about an object (e.g., a movie, a camera, or a car).</li> </ol>	
<b>he Role of the student &amp; benefits gained from participation in this project:<sup>1</sup></b>	<ul style="list-style-type: none"> <li>• Learn about research team environment and work</li> <li>• Practice communication skills</li> <li>• Develop research and programming skills</li> <li>• learn about Opinion Mining and Sentiment Analysis</li> <li>• Publication might be expected</li> </ul>	
<b>Who will be working with you?</b>	Dr. Lamia Tounsi	
<b>Short description of the group:</b>	<p>The National Centre for Language Technology is based in the School of Computing in Dublin city University. The Centre carries out basic and applied research in the areas of machine translation, natural language parsing, grammar induction, question answering, sentiment analysis, computer-aided language learning, software localisation, speech recognition and speech synthesis. Its researchers are drawn from the School of Computing, the School of Applied Languages and Intercultural Studies and the school of Electronic Engineering. The Centre is affiliated with the Center for Next Generation Localisation.</p> <p><a href="http://www.nclt.dcu.ie/">http://www.nclt.dcu.ie/</a></p>	
<b>Recommended Reading Material:</b>	<ol style="list-style-type: none"> <li>1) Mining Arabic Business Reviews, Mohamed Elhawary, Mohamed Elfeky, 2010 IEEE International Conference on Data Mining Workshops, pp. 1108-1113</li> <li>2) Sentiment classification of movie reviews using linguistic parsing, Brian Eriksson, 2008 <a href="http://www.cs.wisc.edu/~apirak/cs/cs838/eriksson_final.pdf">http://www.cs.wisc.edu/~apirak/cs/cs838/eriksson_final.pdf</a>.</li> </ol>	
<b>Other information:</b>		
<b>For further details on this project please contact:</b>	<b>Name:</b> <b>Phone:</b> <b>E-Mail:</b> <b>Website:</b>	<b>Dr. Lamia Tounsi</b> <b>(0)1 700 6905</b> <a href="mailto:lamia.tounsi@computing.dcu.ie">lamia.tounsi@computing.dcu.ie</a>

<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.