CNGL Undergraduate Students as Researchers Programme 2011
PROJECT DESCRIPTION

Institution/Track:	Dublin City University
Project Title:	Experimental Retrieval Engine for Desktop Search
Suitable for students who are studying in the following areas:	Suitable for students who are studying Computing, Computer Science or related disciplines with knowledge of Java.
Skills needed:	Suitable candidates need to be comfortable with Java programming and keen to advance their programming skills. An interest in search technologies, information retrieval and document processing in general is desirable. Knowledge in development of plug and play standalone applications would also be desirable.
Project Description:	Lifelogging is a new and exciting area of research. A lifelog (personal archive) may consist of emails sent or received, web pages visited, files on one's PC or mobile phone, photographs taken, music heard, etc. Our research group is involved in developing techniques to help people search for information in their personal archives. Further progress of our research requires the development of a software tool to collect the files on peoples' PCs and information about these files (e.g., the date a file was last modified, the extension type of the file). The tool should also allow people to type queries and present them with documents which might match their queries. The user should then be able to mark the documents which are of interest to their query. This information will be logged on peoples' PCs to facilitate our experimentation. The intern will work on the development of this local collection and search tool. To do this they will extend components of the 'Terrier' (http://terrier.org/) open source information retrieval research framework. Terrier is written in Java. *Please note: this project is available for 12 weeks.
The Role of the student & benefits gained from participation in this project:[1]	The student will write Java code, with guidance from the PIs, to extend the existing Terrier desktop indexing and search tool to facilitate lifelogging experimentation as described in the project description. The student will gain the opportunity to advance their Java programming skills, learn about empirical evaluations in search, first hand experience of working within a dynamic research group, and knowledge in a cutting edge research area. Depending on the project outcomes it is also foreseen that the project will result in a research publication, of which the student would be a co-author.
Who will be working with you?	The student will work closely with Liadh Kelly, a researcher in the group. S/he will also be supervised by Dr Gareth Jones, a Principal Investigator (PI) in the group.
Short description of the group:	We have 3 people in our personal information access team – Dr Gareth Jones (the team leader/PI) and 2 PhD students.
Recommended Reading Material:	Information retrieval textbook, e.g., Introduction to Information Retrieval by Manning, Schutze and Raghavan. http://terrier.org/ (The Java indexing and search library which will be used in the project).

[1] ***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.***

	A Strategy for Evaluating Search of 'Real' Personal Information Archives. Jones G.J.F. and Chen Y. In *Proceedings of Workshop on Evaluating Personal Search at ECIR 2011*, Dublin, Ireland, 18 April 2011. (The 'living lab' notion described in this paper forms the basis of the exploratory study which this project deals with). The Information Retrieval Challenge of Human Digital Memories. Kelly L. *BCS IRSG Symposium: Future Directions in Information Access 2007*, Glasgow, Scotland, 28-29 August 2007.	
Other information:		
For further details on this project please contact:	**Name:** **Phone:** **E-Mail:** **Website:**	**Gareth Jones** **00 353 1 700 5559** **gjones@computing.dcu.ie** http://computing.dcu.ie/~gjones/