



## Final project proposal

<b>Type</b>	Master
<b>Title</b>	<i>Information Extraction from the WEB</i>
<b>Supervisor</b>	Prof. Igor Trajkovski
<b>e-mail</b>	<a href="mailto:trajkovski@finki.ukim.mk">trajkovski@finki.ukim.mk</a>
<b>Department / Group</b>	
Intelligent Systems	
<b>Topic(s)</b>	
Natural Language Processing, Machine Learning	
<b>Project can start from</b>	01.04.2014
<b>Project duration</b>	4 months
<b>Short description</b>	
<p>Information Extraction is the process of retrieving structured information from unstructured text. Web Information Extraction (WIE) systems differ significantly from traditional systems in both methods and goals. Whereas traditional IE systems focus on squeezing as much juice as possible from small corpora, WIE systems focus on domain independent extraction from relatively simple sentences, and rely on the redundancy of the Web to provide large quantities of information. This thesis should explore a ways in which structured information (entities, relations, etc.) can be extracted from unstructured text.</p>	
<b>Results and assessment</b>	
<p>The outcome of this project should be a written report and oral presentation. The elements of the report should include: introduction, review of literature, procedures used for collecting and analyzing data, findings, summary and conclusion.</p>	
<b>Other (additional) information</b>	
<p><a href="http://www.dfki.de/~neumann/esslli04/reader/overview/IJCAI99.pdf">http://www.dfki.de/~neumann/esslli04/reader/overview/IJCAI99.pdf</a> <a href="http://www.time.mk/trajkovski/">http://www.time.mk/trajkovski/</a></p>	