



Title	Arabic Tokenizer
Description	<p>- It's a part of a search engine which takes a paragraph of Arabic Text and outputs a list the significant <i>Tokens</i> (Words) and removing all the Stop words (unused words) like (حروف الجر والوصل... إلخ).</p> <p>Note: Choose the suitable data structure that helps a faster search</p>
Group size	4-5 members.
Deliverables	<ol style="list-style-type: none"> 1. Creating a Stop Words List. 2. Editing the Stop Words List . 3. Converting the input Text into a Tokens list excluding Stop words. 4. Keeping track of the <i>frequency</i> of repeated tokens.
Bonus extensions	<ul style="list-style-type: none"> - Searching for a Token by getting its position in the list if it exists. - Updating the Tokens list when typing new text. - Any additional module which will help in searching. - Visualizing the Tokens list.
Mentor	A.L Sally S.Ismail
Notes	<p>-You have to check the search engines list and you will understand more.</p> <p>-Check http://en.wikipedia.org/wiki/Web_search_engine for more info</p>