

RK-A - Feature #1809

Checking the context of the KP matched with "word1.word2" header variant

10/27/2021 10:30 AM - Nandini Bansal

Status:	Resolved	Start date:	10/11/2021
Priority:	Normal	Due date:	
Assignee:		% Done:	28%
Category:		Estimated time:	3.00 hours
Target version:	P1	Spent time:	0.00 hour
Description			
<p>We have generally observed if the KP is matching with the "word1.word2" header variant with KP being equivalent to "word2", the linking of KP and the similar doc is out of context at most of the places as these subsection headers are quite specific to the documents they are part of.</p> <p>For e.g. KP = method header variant1 = class.method header variant2 = request.method</p> <p>KP is linked to both "header variant1" and "header variant2" with high confidence scores but they might or might not be relevant as "header variant2" is very specific about the topic it is relevant for.</p> <p>To test this idea, we can fetch the sentence from which the KP is extracted and check if "word1" is present in it. If "word1" is an abbreviation, check for both "word1" & full-form of the "word1" in the context sentence.</p> <p>Added screenshot of KP "dump" getting matched to "pickle.dump", "ast.dump", etc subsections with high confidence but they are out of context due to their specific nature.</p> <p>Implementation in generate_candidates function.</p> <p>Very important to check with the C-API book & Library Reference book.</p>			
Subtasks:			
Feature # 1810: Implement the penalty algorithm for KPs matching with "word1.word2" hea...			Resolved
Feature # 1811: For matching "word1" in with the KP, make use of token_processing function			Resolved
Feature # 1812: Analysis for improvising the algorithms			Rejected
Feature # 1813: Analysis for improvising the algorithms			In Progress

History

#1 - 10/27/2021 10:46 AM - Nandini Bansal

- Status changed from New to In Progress

#2 - 10/28/2021 06:55 AM - Anonymous

- Status changed from In Progress to Resolved

Files

Screenshot from 2021-10-27 15-57-04.png

127 KB

10/27/2021

Nandini Bansal