RK-A - Bug #1507

Bad looking key phrases following the pattern: "word1, word2" while the header variant is "word1 word2"

08/18/2021 03:02 PM - Nandini Bansal

Status:	Closed		**Start date:**	08/18/2021
Priority:	Normal		**Due date:**	
Assignee:			**% Done:**	0%
Category:			**Estimated time:**	2.50 hours
Target version:			**Spent time:**	0.00 hour

Description

In Python Whirlwind Tour.txt, we are seeing some tagged KPs such as "python, code" linked to the "How to run Python Code" document. It is certainly undesirable. Some of such instances were observed in Python 3 - Library Reference.txt as well.

Make changes in BR3_IR3_tagger.py to make sure such KPs are eliminated. Test the changes with the following datasets:
1. Python Whirlwind Tour.txt
2. Python Tutorial.txt
3. Python 3 - Library Reference.txt

For running "Python 3 - Library Reference.txt", you can use master_cands_0.pkl from this link - https://drive.google.com/file/d/1QdZvcVPnPHR4XfKRwscs8PWBJGskRYcF/view?usp=sharing

History

#1 - 08/19/2021 06:49 AM - Nandini Bansal

In Python Whirlwind Tour.txt, we are seeing some tagged KPs such as "python, code" linked to the "How to run Python Code" document. It is certainly undesirable. Some of such instances were observed in Python 3 - Library Reference.txt as well.

Make changes in BR3_IR3_tagger.py to make sure such KPs are eliminated. Test the changes with the following datasets:
1. Python Whirlwind Tour.txt
2. Python Tutorial.txt
3. Python 3 - Library Reference.txt

For running "Python 3 - Library Reference.txt", you can use master_cands_0.pkl from this link - https://drive.google.com/file/d/1QdZvcVPnPHR4XfKRwscs8PWBJGskRYcF/view?usp=sharing

URL:
https://edutestdev-240612.appspot.com/document/python-whirlwind-tour/m?documentURL=10054%2Fds9aug1528%2FWhirlwindTourOfPython%2F02-basic-python-syntax-Indentation-Whitespace-Matters!-13.html

#2 - 10/13/2021 12:35 PM - Anonymous

- Status changed from New to Resolved

#3 - 10/20/2021 08:35 AM - Nandini Bansal

- Status changed from Resolved to Closed

Files

Screenshot from 2021-08-18 20-27-27.png	68.4 KB	08/18/2021	Nandini Bansal