

RK-A2 - Feature #3642

Task # 3653 (Closed): Launch with first chapter support - Phase 1 - Practice for Decent Learners (Top 5 Themes, No DetailedQNAs, No Detailed Answers)

Changes to call_openai_api to handle context limitations

07/01/2024 06:49 AM - Ram Kordale

Status:	Closed	Start date:	07/01/2024
Priority:	Normal	Due date:	
Assignee:	Ram Kordale	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:		Spent time:	0.00 hour
Description			
Add an additional argument num_tokens. Add an additional output num_tokens.			
Clients interested in handling the context limitation must send num_tokens received in the previous call in the thread as argument. First time, the value is 0.			
As always, get assistant, instructions, prompt_prefix and prompt_suffix based on model and index. In addition, also get periodic_prompt from DS.			
So far, prompt used is concat(prompt_prefix, <prompt>, prompt_suffix). If num_tokens is not zero, continue this. Otherwise prompt should be concat(periodic_prompt, prompt_prefix, <prompt>, prompt_suffix).			
One more change: add response['usage']['total_tokens'] (tokens used in this api call) to num_tokens and return it.			

History

#1 - 07/01/2024 09:45 AM - Md Shahzar

- Status changed from New to In Progress

#2 - 07/01/2024 10:17 AM - Md Shahzar

- Status changed from In Progress to Resolved

- % Done changed from 0 to 100

#3 - 07/06/2024 12:40 PM - Ram Kordale

- Parent task set to #3653

#4 - 07/09/2024 06:56 AM - Md Shahzar

- Status changed from Resolved to Feedback

- Assignee changed from Md Shahzar to Ram Kordale

#5 - 09/11/2024 06:24 AM - Ram Kordale

- Status changed from Feedback to Closed