

# Retriever:

- 1) adding tool descriptions and integrating retriever:  
<https://colab.research.google.com/drive/1Bbm9LGU82tSLPqVFXEcLIL7facqnmGr2?usp=sharing>
- 2) (trainable llm file)IR bert training and Mini lm v2 training:  
<https://colab.research.google.com/drive/15-0ulzqlJOTkv5cmNENwpBXMcuXDiBeJ?usp=sharing>
- 3) OpenAI Retriever:  
<https://colab.research.google.com/drive/1d5Bu66srTZy-NUGNPx7J0YbwRNx6plaq?usp=sharing>
- 4) BM25, Contriever, mini LM L6 v2:  
[https://drive.google.com/file/d/1hwIDYTH\\_Yi7K6RnU87jNljSUi0mPk3bN/view?usp=sharing](https://drive.google.com/file/d/1hwIDYTH_Yi7K6RnU87jNljSUi0mPk3bN/view?usp=sharing)

# Data generation:

- 1) thought-action pair generation code=>  
[https://colab.research.google.com/drive/1jh9l6kFO76fsTG\\_5q7BSWbuV4mPiR1\\_c?usp=sharing](https://colab.research.google.com/drive/1jh9l6kFO76fsTG_5q7BSWbuV4mPiR1_c?usp=sharing)
- 2) query-relevant api dataset generation (input data for thought action pair)=>  
<https://colab.research.google.com/drive/1PF19iIB65YOv4zeNhwWzErgP0MhrGR-X?usp=sharing>
- 3) query-json response (direct) data generation=>  
<https://colab.research.google.com/drive/1huFozlBdYV2p1prgN6bpkDjQklRuQKyb?usp=sharing>

# Train code:

- 1) training\_ToolExecutor\_AlternateApproach1\_step4\_option3=>  
<https://colab.research.google.com/drive/1k8Zs53Q5MCgaaE6qoodR7LqbODqjMU4W?usp=sharing>
- 2) JsonFormer=>  
<https://colab.research.google.com/drive/1fzIZsuFLJtrLxNhbehu0ZZULjWXsvgL#scrollTo=phozRyyfwd9E>
- 3) Final code for training llms -> using lora -> training can run on google colab=>  
[https://colab.research.google.com/drive/1T-RPScdS5l3bbp0nOtBbS05uiOM\\_yslQ?usp=sharing](https://colab.research.google.com/drive/1T-RPScdS5l3bbp0nOtBbS05uiOM_yslQ?usp=sharing)

# Deployment:

- 1) run chainlit application from google colab=>  
<https://colab.research.google.com/drive/1lufjSR2E4wXCrWriLdzOGF8pR4oRt8QW?usp=sharing>
- 2) convert to gglm and Quantize model =>  
[https://colab.research.google.com/drive/1NSPIlh3o9apPuY8UdpSUZQYOTVgG\\_t\\_0?usp=sharing](https://colab.research.google.com/drive/1NSPIlh3o9apPuY8UdpSUZQYOTVgG_t_0?usp=sharing)
- 3) Model deployment on chainlit =>  
<https://drive.google.com/file/d/1pbMYFDCvQsdKP294WZgtj8HSsGFTWFCC/view?usp=sharing>

# Bonus:

- 1) running inference on Microsoft TORA =>  
<https://colab.research.google.com/drive/16MuwhMWPnBpEaqbUqtuU1UM-QVGkSqeI?usp=sharing>

# Final\_colab:

- 1) Final colab :  
<https://drive.google.com/file/d/1JB7QSBE-Knk0fMZnUqt2zPODRLNf62G2/view?usp=sharing>

# Dataset:

[https://drive.google.com/drive/folders/1\\_ydX-CZQW1OrXtvR7ISp8E2wcaRe-UFS](https://drive.google.com/drive/folders/1_ydX-CZQW1OrXtvR7ISp8E2wcaRe-UFS)