

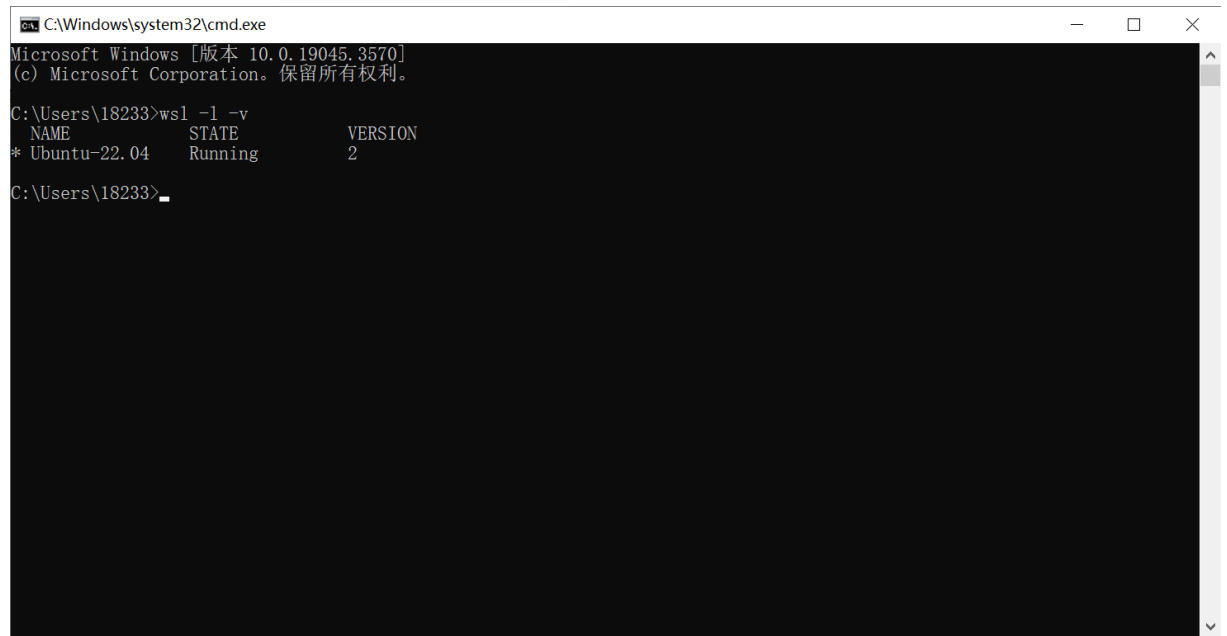
# 运行配置说明

211870287 丁旭

October 2023

## 1 Introduction

### 1. 安装 wsl2



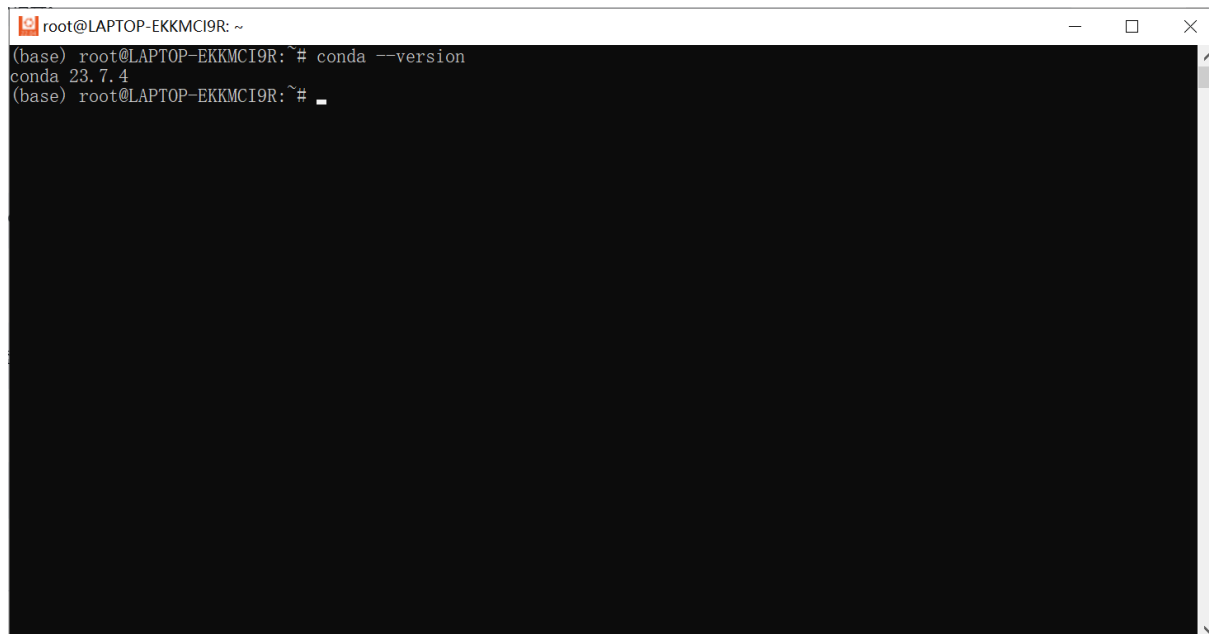
```
C:\Windows\system32\cmd.exe
Microsoft Windows [版本 10.0.19045.3570]
(c) Microsoft Corporation。保留所有权利。

C:\Users\18233>wsl -l -v
  NAME                STATE      VERSION
* Ubuntu-22.04        Running    2

C:\Users\18233>
```

图 1: ubuntu22.04

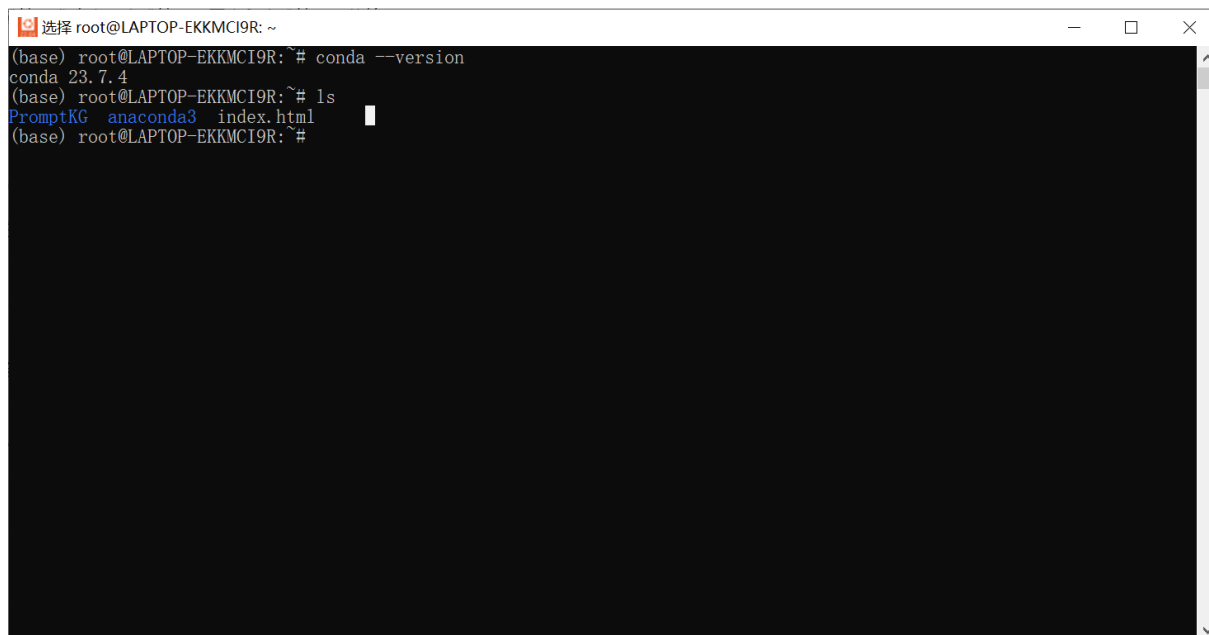
## 2. 安装 annaconda

A terminal window with a black background and white text. The title bar at the top reads 'root@LAPTOP-EKKMC19R: ~'. The terminal content shows a command prompt '(base) root@LAPTOP-EKKMC19R: ~#' followed by the command 'conda --version'. The output of the command is 'conda 23.7.4'. The prompt then returns to '(base) root@LAPTOP-EKKMC19R: ~#'.

```
root@LAPTOP-EKKMC19R: ~  
(base) root@LAPTOP-EKKMC19R: ~# conda --version  
conda 23.7.4  
(base) root@LAPTOP-EKKMC19R: ~#
```

图 2: conda23.7.4

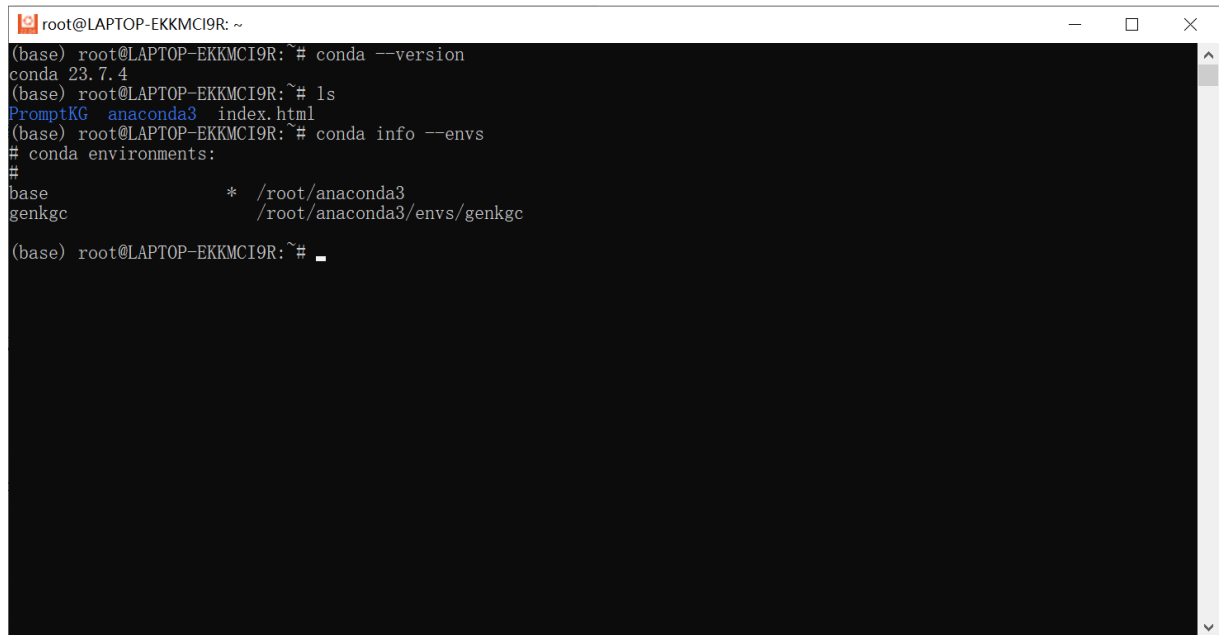
3. 下载源代码 `git clone -depth 1 https://github.com/zjunlp/PromptKG.git`



```
选择 root@LAPTOP-EKKMCI9R: ~
(base) root@LAPTOP-EKKMCI9R: ~ # conda --version
conda 23.7.4
(base) root@LAPTOP-EKKMCI9R: ~ # ls
PromptKG  anaconda3  index.html
(base) root@LAPTOP-EKKMCI9R: ~ #
```

图 3:

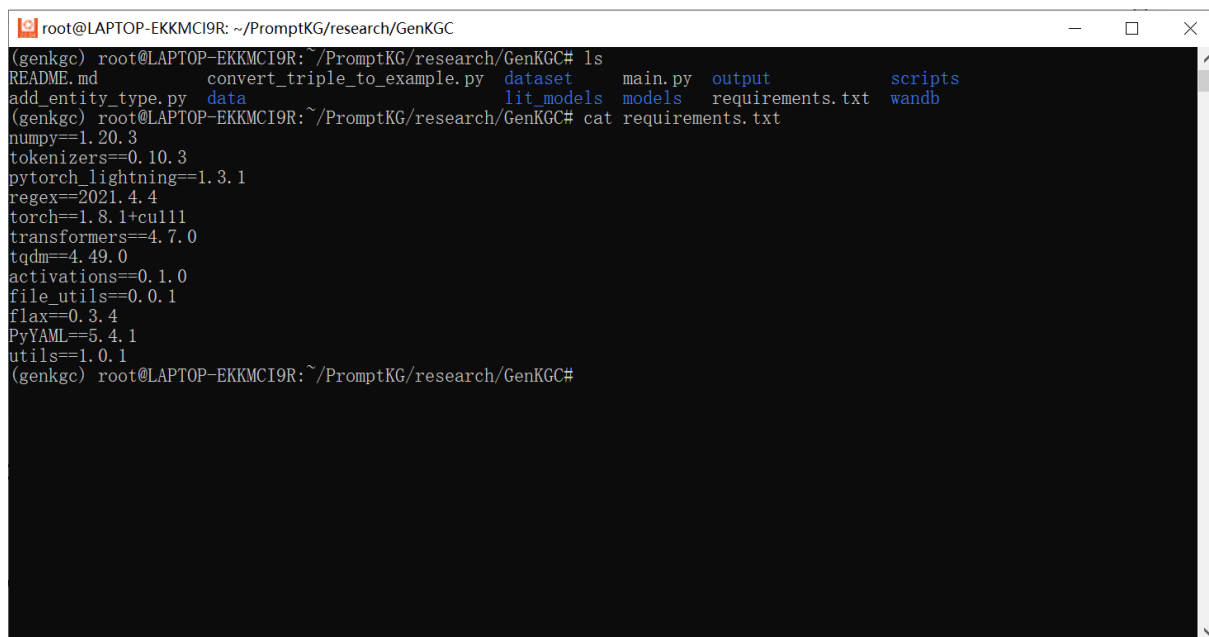
4. 创建虚拟环境 `conda create -n genkgc python=3.8` 并激活 (`conda activate genkgc`)



```
root@LAPTOP-EKKMCI9R: ~  
(base) root@LAPTOP-EKKMCI9R:~# conda --version  
conda 23.7.4  
(base) root@LAPTOP-EKKMCI9R:~# ls  
PromptKG  anaconda3  index.html  
(base) root@LAPTOP-EKKMCI9R:~# conda info --envs  
# conda environments:  
#  
base          * /root/anaconda3  
genkgc        /root/anaconda3/envs/genkgc  
(base) root@LAPTOP-EKKMCI9R:~#
```

图 4:

5. 进入目录 (cd PromptKG/research/GenKGC) 下载依赖 (pip install -r requirements.txt)



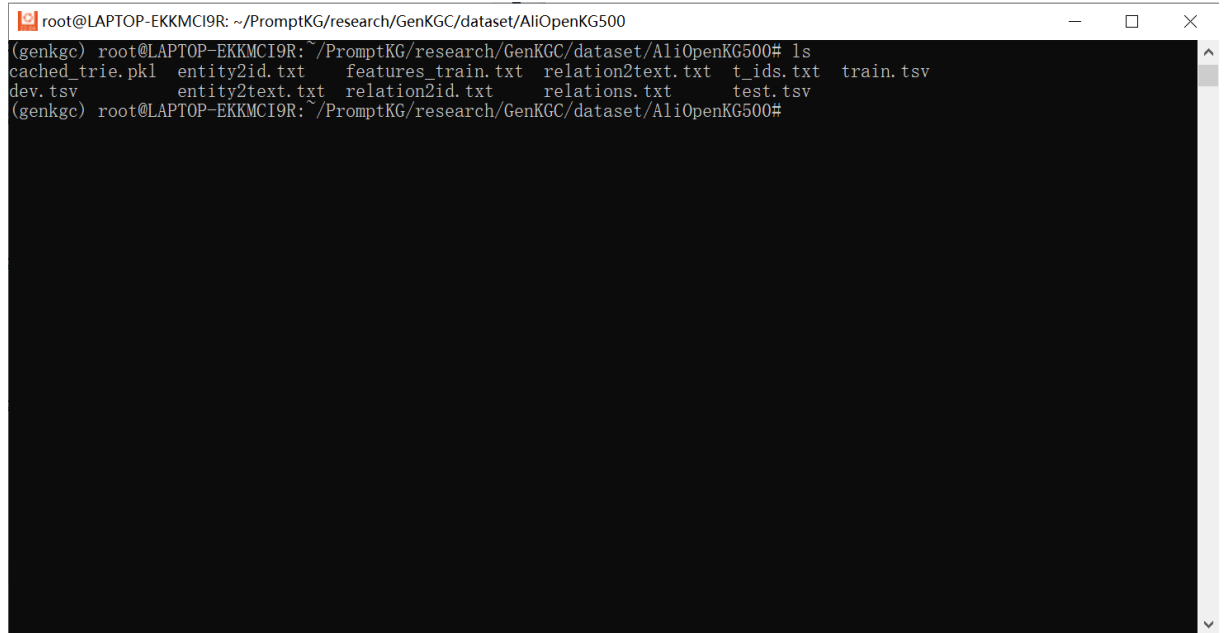
```
root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKGC
(genkgc) root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKGC# ls
README.md      convert_triple_to_example.py  dataset      main.py      output      scripts
add_entity_type.py  data                          lit_models   models       requirements.txt  wandb
(genkgc) root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKGC# cat requirements.txt
numpy==1.20.3
tokenizers==0.10.3
pytorch_lightning==1.3.1
regex==2021.4.4
torch==1.8.1+cu111
transformers==4.7.0
tqdm==4.49.0
activations==0.1.0
file_utils==0.0.1
flax==0.3.4
PyYAML==5.4.1
utils==1.0.1
(genkgc) root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKGC#
```

图 5: cat requirements.txt

6. pip 安装的所有包



7. 下载数据集 git clone <https://github.com/OpenBGBenchmark/OpenBG500.git>



```
root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKG/dataset/AliOpenKG500
(genkgc) root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKG/dataset/AliOpenKG500# ls
cached_trie.pkl  entity2id.txt  features_train.txt  relation2text.txt  t_ids.txt  train.tsv
dev.tsv          entity2text.txt  relation2id.txt     relations.txt      test.tsv
(genkgc) root@LAPTOP-EKKMCI9R: ~/PromptKG/research/GenKG/dataset/AliOpenKG500#
```

图 8:

## 8. 运行代码./scripts/openbg.sh

```
root@LAPTOP-EKMC79R: ~/PromptKG/research/GenKGC
(genkgc) root@LAPTOP-EKMC79R: ~/PromptKG/research/GenKGC# ./scripts/openbg.sh
global seed set to 7
tokenize the prefix tree: 0% | 0/269658 [00:00:?, ?it/s]
化妆喷雾 [101, 8335, 7816, 6857, 23292, 102]
tokenize the prefix tree: 0% | 1/269658 [00:00:8:39:15, 8.66it/s]
葡萄酒 [101, 11072, 21562, 19731, 102]
长袖上衣 [101, 22915, 19684, 4906, 19633, 102]
内裤 [101, 156, 9345, 102]
tokenize the prefix tree: 100%
*****max output length : 128*****
total 0 cannot be tokenized
wandb: (1) Create a W&B account
wandb: (2) Use an existing W&B account
wandb: (3) Don't visualize my results
wandb: Enter your choice: 3
wandb: You chose "Don't visualize my results"
wandb: WARNING resume will be ignored since W&B syncing is set to "offline". Starting a new run with run id afqmeuot.
wandb: Tracking run with wandb version 0.15.12
wandb: W&B syncing is set to "offline" in this directory.
wandb: Run wandb online or set WANDB_MODE=online to enable cloud syncing.
Using native 16bit precision.
CPU available: True, used: True
TPU available: False, using: 0 TPU cores
GPU available: False, using: 0 GPUs
Delete entities without text name.: 100% 5000/5000 [00:00:00:00, 1118361.77it/s]
total entity not in text : 0
max number of filter entities : 18713 2182
convert text to examples: 100% 5000/5000 [00:00:00:00, 280641.80it/s]
100% 5000/5000 [00:01:00:00, 2675.38it/s]
delete entities without text name.: 100% 5000/5000 [00:00:00:00, 1083183.72it/s]
total entity not in text : 0
max number of filter entities : 18713 2182
convert text to examples: 100% 5000/5000 [00:00:00:00, 239450.12it/s]
100% 5000/5000 [00:01:00:00, 2609.81it/s]
LOCAL_RANK: 0 - CUDA_VISIBLE_DEVICES: [0]
/root/anaconda3/envs/genkgc/lib/python3.8/site-packages/torch/utils/data/dataloader.py:474: UserWarning: This DataLoader will create 32 worker processes in total. Our suggested max number of worker in current system is 16, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.
  warnings.warn(_create_warning_msg(
  Name      Type      Params
  -----
0  model     BartKGC
1  loss_fn   CrossEntropyLoss 140 M
140 M      Trainable params
0          Non-trainable params
140 M      Total params
560.778    Total estimated model params size (MB)
Validation sanity check: 0% | 0/2 [00:00:?, ?it/s]
origin input : 无核白葡萄干系列
model output : 无核白葡萄干系列
model constrained output : 无核白葡萄干
无核红葡萄干
无核绿葡萄干
无核葡萄干
无核白
true label : 葡萄干
*****new
global seed set to 7
/root/anaconda3/envs/genkgc/lib/python3.8/site-packages/pytorch_lightning/utilities/data.py:71: UserWarning: Your 'IterableDataset' has '__len__' defined. In combination with multi-process data loading (when num
```

图 9:

## 9. 运行过程和结果



```
root@LAPTOP-EKKMC19R:~/PromptKG/research/GenKGC
login input: 0%无核白葡萄干系列
model output: 无核白葡萄干系列
model constrained output: 无核白葡萄干
无核白葡萄干
无核白葡萄干
无核白葡萄干
无核白葡萄干
无核白葡萄干
true label: 葡萄干
*****
Epoch 2: 2%
login input: 2%脚离志异细分市场
model output: 脚离志异细分市场
model constrained output: 交通工具
图书馆
神话传说
图片
图书
true label: 教材
*****
Epoch 2: 4%
login input: 4%纽曼 MC16 车载麦克风全民唱歌手机 k 歌麦克风唱歌吧话筒蓝牙无线车载 ktv 汽车通用车内唱歌神器直播 1 声卡 主播细分市场
model output: 纽曼 MC16 车载麦克风全民唱歌手机 k 歌麦克风唱歌吧话筒蓝牙无线车载 ktv 汽车通用车内唱歌神器直播 1 声卡 主播细分市场
model constrained output: 北京工艺美术出版社
北京工业大学出版社
北京电子工业出版社
北京大学出版社
美国国家地理
true label: 直播
*****
Epoch 2: 6%
login input: 6%针织套头衫服装款式细节
model output: 针织套头衫服装款式细节
model constrained output: 针织套头衫
针织套头衫
针织连衣裙
针织套头衫
针织套头衫
true label: 字母
*****
Epoch 2: 8%
login input: 8%牛仔裤裤门襟
model output: 牛仔裤裤门襟
model constrained output: 牛仔裤
牛仔裤子
牛仔衣裤
服装
true label: 拉链
*****
Epoch 2: 10%
login input: 10%t 恤服装款式
model output: t 恤服装款式
model constrained output: t 恤
t 恤
t 恤套装
西服套装
西服外套
true label: 宽松
*****
Epoch 2: 12%
login input: 12%女黑 t 恤服装
model output: 女黑 t 恤服装
model constrained output: 女黑 t 恤服装
```

图 10:

```
wandb: Run history:
wandb: Eval/hits1
wandb: Eval/hits10
wandb: Eval/hits20
wandb: Eval/hits3
wandb: epoch
wandb: trainer/global_step
wandb:
wandb: Run summary:
wandb: Eval/hits1 0.2032
wandb: Eval/hits10 0.3684
wandb: Eval/hits20 0.4844
wandb: Eval/hits3 0.2804
wandb: epoch 8
wandb: trainer/global_step 8
wandb:
wandb: You can sync this run to the cloud by running:
wandb: wandb sync /root/PromptKG/research/GenKGC/wandb/offline-run-20231013_170516-afqmwuek
wandb: Find logs at: ./wandb/offline-run-20231013_170516-afqmwuek/logs
(genkgc) root@LAPTOP-EKKMC19R:~/PromptKG/research/GenKGC#
```

图 11: