JK17K is a subset of the open-domain knowledge graph named FREEBASE.

Firstly, entities involved in very few triples and the triples involving String, Enumeration Type and Numbers were removed. Then, a fact representation was recovered from the remaining triples. Facts from meta-relations having only a single role were also removed. Next, we extract the most frequent meta-relations and their facts, and convert them into instance representation. Finally, the subset was divided into training/testing sets randomly, ensuring each entity in test set appeared in training set. Note that, for those instances belonging to the same fact, there is also at least one instance was placed in the training set.

There are totally three version of JF17K applicable for different model, obviously, they are equivalent.

## Version1: Instance:

Each line represents one instance. The first column is the index of instance. The second column is the type of multifold relation, the reset are mid of entities belonging to this instance ordered by the schema of this relation.

Entity.txt: entities

Relation.txt: relation + tab + number of folder(N)

Train.txt: tab + type of relation + entity1 + entity2 + … + entityN

Test.txt: instance\_index + tab + type of relation + head entity + tail entity

## Version2: Instance with dummy node:

Each line represents one instance. The first column is the index of instance. The second column is the type of multifold relation; for those relations which need dummy node, the third column is the dummy node’s mid (cvt-id); the reset are mid of entities belonging to this instance ordered by the schema of this relation.

Entity.txt: entities and cvt-ids

Relation.txt: relation + tab + number of folder(N)

Real entity: entity.txt minus cvt-ids

Train.txt: tab + type of relation + entity1 + entity2 + … + entityN

Test.txt: instance\_index + tab + type of relation + head entity + tail entity

## Version3: Triple:

After applying star-to-clique convert on Version1, we obtain the third Version. Each line represent a triple.

Entity.txt: one entity each line

Relation.txt: relation + tab + number of folder

Train.txt: tab + type of relation + head entity + tail entity

Test.txt: instance\_index + tab + type of relation + head entity + tail entity