各位师弟：

大家好，师兄需要你们帮忙做一个数据的评价，希望大家帮帮忙。罗老师已经同意。

目前有两个文档，一个entity.txt文档，一个relation.txt文档,entity.txt文档中包含有100个实体对，relation.txt文档中包含有大约8000个关系。

具体要求如下：

1. entity文档中的实体对进行评价标准：只要该实体对有关系，就认定该实体对是准确的。怎样才算有关系？可以将一个实体对中的两个实体在百度查询，比如：vitreous hemorrhage（玻璃体积血）,retinal detachment（视网膜脱离）这个实体对，因为百度查询结果显示，玻璃体积血会引起视网膜脱离，那么这个实体对就是有关系的，这个实体对就是准确的。

四个人将entity文档中100个实体对均分一下，每人25个，最后将每个统计的结果合并，给出一个最终的准确率。将准确的实体对在word中用颜色标注出来，比如：

vitreous hemorrhage（玻璃体积血）,retinal detachment（视网膜脱离）

pressure,glaucoma

glaucoma,osteoporosis（骨质疏松）

其中红色表示准确，无色表示错误，（最好将每个英文实体的意思写出来）。

2. 对于每个实体对，在relation文档中都有对应的关系表述，大家对每个实体对所有的关系描述随机挑选四个，判断这四个关系是否准确（最终每个人会有一百个关系需要去判断）。

怎么样才算关系准确？比如：

{<pressure>,<in primary open-angle >,<glaucoma>，描述的意思大概是：压力在在开放性青光眼中存在。这就算是准确的。

{<pressure>,<is a well-known major risk factorfor >,<glaucoma>，压力是青光眼主要的危险因子。

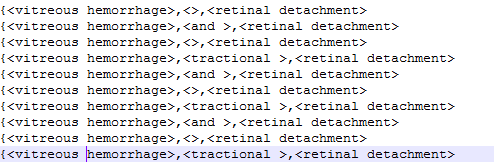
这也是准确的。

{<vitreous hemorrhage>,<and >,<retinal detachment>,玻璃体积血和视网膜脱离，这个没有实际意义，目前可以认为是错误的。

{<vitreous hemorrhage>,<tractional >,<retinal detachment>，玻璃体积血牵引出视网膜脱离。

这是准确的。

至于relation中重复的语句，比如：



按不重复处理。

每个人将各自的100个关系进行准确性计算，最后整合到一个word文档，将准确的关系用颜色标注。再算出一个总体的准确性。

谢谢各位师弟，希望大家能认真对待这次评价，并且客观评价。给我一个准确的反馈结果，辛苦大家了。拜托了！如有什么不懂或不清楚的地方，请大家及时提问。感激不尽！