# 功能描述

1. 基于DBpedia链接开放数据实现英文出版物(电影、图书、游戏) 检索和推荐系统。
2. 实现出版物的检索功能，能够直接检索到相关出版物的基本摘要，并给出相关介绍网页的Wikipedia链接。
3. 实现出版物语义推荐的功能，基于知识图谱给出的语义，推荐用户可能感兴趣的其他出版物。第一期思路主要是基于具有共同主题的实体进行推荐。
4. 实现对出版物相关数据的可视化分析与展示。
5. 所有源码上传到GitHub上：<https://github.com/LongXinchen/kg_recommendation> 。

# 设计思路

## 检索功能的设计

1. 具体需求
   1. 输入出版物(电影、图书、游戏)的（大小写敏感）英文名称query，返回名称包含该query的实体、Wikipedia链接和英文摘要。
2. 设计思路
   1. 采用正则匹配的方法，构造基于dbpedia的sparql，返回类型属于(电影、图书、游戏)的实体的相关字段。
3. sparql如下：

'''

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>

PREFIX dbo:<http://dbpedia.org/ontology/>

PREFIX dc:<http://purl.org/dc/terms/>

SELECT distinct ?name ?url ?url2 ?abstract ?work

WHERE {

{?work rdf:type dbo:Book ;

rdfs:label ?name ;

foaf:isPrimaryTopicOf ?url;

prov:wasDerivedFrom ?url2;

dbo:abstract ?abstract

FILTER(lang(?name)="en" &&(lang(?abstract)="en") && regex(str(?name),"%s")).

}

UNION

{?work rdf:type dbo:Film ;

rdfs:label ?name ;

foaf:isPrimaryTopicOf ?url;

prov:wasDerivedFrom ?url2;

dbo:abstract ?abstract

FILTER(lang(?name)="en" &&(lang(?abstract)="en") && regex(str(?name),"%s")).

}

UNION

{?work rdf:type dbo:VideoGame ;

rdfs:label ?name ;

foaf:isPrimaryTopicOf ?url;

prov:wasDerivedFrom ?url2;

dbo:abstract ?abstract

FILTER(lang(?name)="en" &&(lang(?abstract)="en") && regex(str(?name),"%s")).

}

}

''' % (query, query, query)

## 推荐功能的设计

1. 具体需求
   1. 输入出版物(电影、图书、游戏)的（大小写敏感）英文名称query，返回TOP10共享最多类别的实体的Wikipedia链接。
2. 设计思路
   1. 基于检索得到的实体，构造基于dbpedia的sparql，返回共享最多类别的实体对应的Wikipedia链接。
3. sparql如下：

'''

PREFIX rdf:<http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>

PREFIX dbo:<http://dbpedia.org/ontology/>

PREFIX dct:<http://purl.org/dc/terms/>

SELECT COUNT(?work) SAMPLE(?work)

WHERE

{

dbr:%s dct:subject ?o .

?work dct:subject ?o

FILTER (?work != dbr:%s ) .

}

GROUP BY ?work

ORDER BY DESC(COUNT(?work))

limit 10

''' % (url, url)

## 结果可视化的设计

1. 具体需求
   1. 检索与推荐的结果均采用html网页进行呈现，输出文件为output.html 。
   2. 对于检索得到的结果，采用“Search”作为二级标题，分列表展示对应实体的标题、Wikipedia链接、摘要。
   3. 对于推荐得到的结果，采用“Recommend”作为二级标题，分列表展示对应实体的标题、Wikipedia链接。