北京航空航天大学研究生课程考核记录

2017－2018学年 第二学期

学号 姓名 马健

学号 姓名 殷晓婷

学号 姓名 刘子晗

学号 姓名 雷一明

课程名称：《 知识工程基础 》

题目：**基于知识图谱的出版物检索和推荐系统**

日期 ：2018年6月20日

**目录**

1.概述……………………………………………………………………………………………1

2.设计思路………………………………………………………………………………………2

2.1项目需求……………………………………………………………………………2

2.2确定系统的运行环境…………………………………………………………………2

2.3划分模块………………………………………………………………………………2

2.4实现功能及测试……………………………………………………………………2

3.功能描述………………………………………………………………………………………3

4.使用说明………………………………………………………………………………………4

4.1系统检索功能……………………………………………………………………4

4.2系统检索统计功能………………………………………………………………5

4.3系统推荐功能………………………………………………………………………6

5.总结………………………………………………………………………………………

1. **概述**

近年来人工智能技术发展迅速，由此带动了知识图谱等新兴学科的发展。当前，知识图谱在国内外得到了一定的研究和应用，而信息推荐技术已经被广泛应用到互联网以及移动互联网产品中。今日头条，UC浏览器等应用都实现了对用户的精准个性化推荐。电影，图书，游戏等出版物的搜索与推荐成为新媒体公司的重要工作。我们选取知识图谱和信息推荐技术作为支撑点，并从实现技术层面对基于知识图谱的出版物检索和推荐系统个性化推荐方法开展研究。对比分析了当前流行的几种信息推荐技术，选择了基于用户和产品的协同过滤这一个性化推荐技术，同时结合知识图谱理论及其它相关技术，构建了一个基于知识图谱的出版物检索和推荐系统，给出了系统实现结果。

1. **设计思路**

**2.1 项目需求**

首先我们根据作业要求，了解要实现的功能，需要注意的事项及技术难点。需要用到HTML、CSS、Javascript等知识。

* 1. **确定系统的运行环境**

我们系统是在window系统上用谷歌浏览器打开，edge浏览器打开可能会出现一些bug。

* 1. **划分模块**

根据功能需求，划分为搜索，推荐，可视化等三大模块。先实现每一个单独模块，然后再组合在一起

* 1. **实现功能及测试**

完成需求所要实现的功能，然后不断测试修改bug。

**3.功能描述**

我们的基于知识图谱的出版物检索和推荐系统能够搜索到的出版物包括：电影、图书和游戏三类。利用DBpedia数据，实现了出版物的检索并展现出来，还给出了相关介绍网页的链接。在搜索结果中基于知识图谱给出的语义，推荐了一些用户可能感兴趣的其他出版物，并实现了对出版物相关数据的可视化分析与展示，包括检索结果中部分数据的可视化展示以及对于出版物的统计分析等。

**4.使用说明**

**4．1系统检索功能**

基于知识图谱的出版物检索和推荐系统整体包括三个部分的功能：检索、推荐以及可视化。首先，在检索时可以选择搜索出版物的种类，通过点击四个单选框的按钮确定自己要搜索的出版物类别，电影、图书、游戏或者不限类别，然后在检索框中输入待检索信息，点击search按钮即可完成搜索过程。检索过程示例如下图所示。



**图1 系统检索示意图**

当分别选择电影、图书、游戏三个出版物类别时，系统会将对应类别的搜索结果信息显示出来，在界面上显示的信息检索结果包括出版物的图片、语言、国家、作者、相关链接以及出版物的简介。检索结果如下图所示。



**图2 系统检索结果展示图**

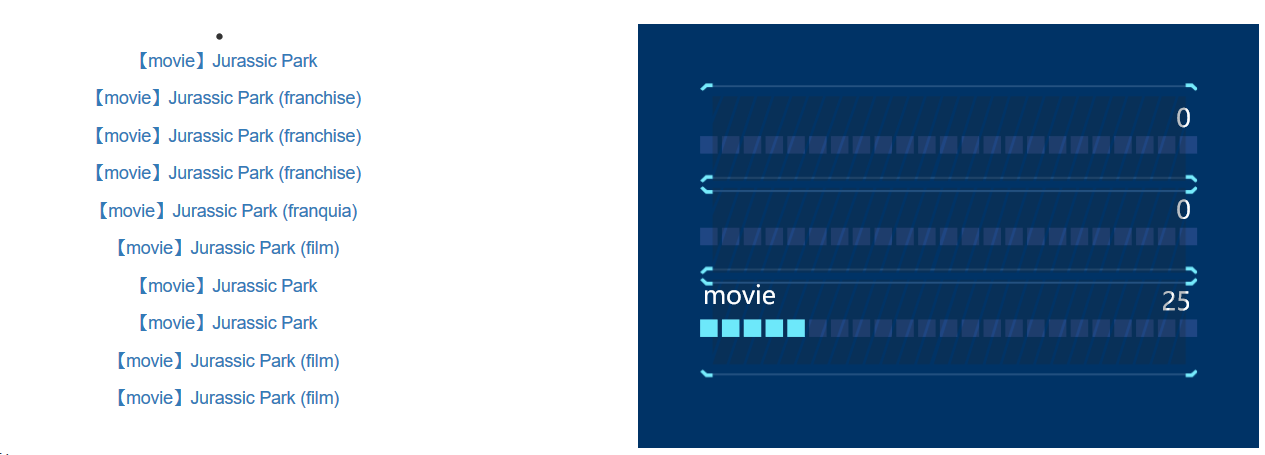
其中，如果结果中包含语言、国家、作者的信息，可以点击相应的信息跳转到其对应的维基百科的结果。



**图3 系统结果链接结果图**

**4．2系统检索统计功能**

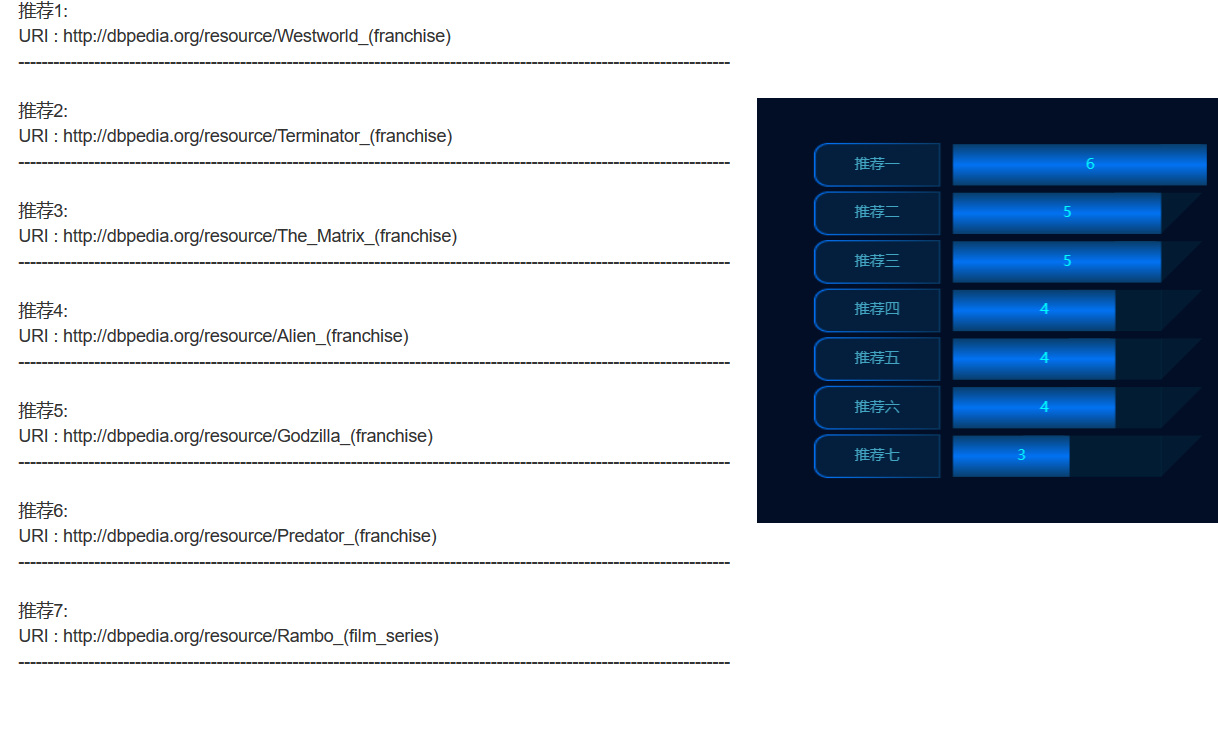
在检索时，除了检索结果信息显示之外，还有对检索结果的简单统计。统计结果是一个包含有出版物的类别、名称和国家的列表，此列表只显示10条结果信息，在列表的右侧对应的是对各个类别出版物的检索结果数量的统计，统计结果显示是三个进度条的形式，三个进度条分别对应电影、图书以及游戏，下方分别显示对应的类别以及出版物检索结果的数量。比如在检索Jurassic Park时，检索结果包含25个不同的电影实体，其中十条结果显示在左侧列表中，显示结果的格式是“[出版物类别]出版物名称(出版物国家)”，右侧的可视化结果显示搜索该实体的结果包含25条电影类型的结果。



**图4 系统检索统计结果图**

**4．3 系统推荐功能**

在检索的同时，系统会对关键字进行相关的语义推荐，推荐结果列表是根据与该检索关键字类别相似度得到的，两个实体在系统中同时出现的次数越多，推荐系数越高。其中推荐列表给出了其推荐实体结果对应的数据库链接，还给出了每个推荐结果对应的推荐因子指数，按照推荐指数的高低进行排序，推荐列表和可视化结果如下图所示。



**图5 系统推荐结果示意图**

**5.总结**

项目开始时队长马健让大家一起收集资料，然后根据作业要求设计基于知识图谱的出版物检索和推荐系统。系统设计完毕后就由马健，殷晓婷，刘子晗完成系统的检索，推荐，可视化等功能。最后由雷一明根据我们制作的系统完成相关文档的撰写。根据各人能力特长知人善用，能够较好的完成一个团队项目。

我们根据项目计划有一个大概的时间进度安排。人员，时间都安排的还算合理。根据最初要求需要6月18日交作业，我们那几天加班加点找资料，向大牛请教问题。由于我们都没有这种项目经验，不求把系统完成很完美，只求保证把作业要求的功能实现。后来作业提交时间延期一个星期后，我们又在之前做的项目进行改善，以求达到较好的效果。最后我们的系统不仅可以为用户呈现较为完整的检索内容 还预测了用户可能感兴趣的其他出版物。

通过这个课程，我们学到了知识的获取，表示和使用的方法。它对我们来说是一个全新的领域，我们可以用所学的东西，作为以后工作中研制各类智能系统的基本工具。对很多人来说算是第一次接触这个这个新兴领域，很多概念不是很清楚了解，特别是实践起来有点困难。希望老师在以后的课程中能够多详细地讲解一些例子，然后根据这些例子布置作业，在实践中教学会使学生学的更快更好。