**基于知识图谱的问答系统功能介绍**

**1. 原理**  
  KGQA简单讲就是将问题带入提前准备好的知识图谱寻求答案的一种基于知识图谱的问答系统。该问答系统可以解析输入的自然语言问句，主要运用分词工具和正则表达式匹配得到结果，然后利用对应的Cypher查询语句，请求后台基于Neo4j知识图谱数据库的服务，最终得到我们想要的结果。

**2. 流程**  
  1.实体检测，获取问题的关键词，比如问题“李白写了哪些诗？”，那么首先必须找到李白，才可以进行下一步。  
  2.目的获取，一个问题，我们只获取了实体还不够，比如上面，只有李白，还要有目的，不然可能我是想问李白是哪个朝代的人，哪里的人等等，所以需要找到问题的真实目的。  
  3.关系预测，有了实体和目的，那么我们就需要在知识图谱里面寻找双方的关系，想办法联系起来。  
  4.查询构建，将处理好的三元组带入知识图谱搜索答案。

**3. 实例分析**  
**3.1 数据准备**

本文基于网络爬虫获取了 108 854 条军事武器信息， 主要包括飞行器和舰艇等 8大类武器、 轰炸机等 100余小类， 共计 5 858项武器实体和 184类实体属性类。采集的军事武器信息以实体⁃属性⁃属性值和实体⁃关系⁃实体的方式存储在图数据库 Neo4j中。

**3.2 数据建模**

本文采用 Python 中的 py2neo 模块将数据导入图数据库 Neo4j中， 实现对军事武器知识图谱的存储和可视化展示， 使军事武器领域知识具有更强的可读性， 且方便用户快速获取及理解军事装备领域知识及知识结构。图 4是军事武器知识图谱的部分实例， 图中不同颜色的节点表明不同类型实体， 节点间的边描述了实体间关系。

**3.3 查询实践**

数据建模完成后，我们可以类似使用SQL查询关系数据库一样使用Cypher查询Neo4j数据库中的数据。  
  Cypher是一个描述性的图形查询语言，允许不必编写图形结构的遍历代码对图形存储有表现力和效率的查询。Cypher还在继续发展和成熟，这也就意味着有可能会出现语法的变化。同时也意味着作为组件没有经历严格的性能测试。这个查询语言包含以下几个明显的部分：

START：在图中的开始点，通过元素的ID或所以查找获得。

MATCH：图形的匹配模式，束缚于开始点。

WHERE：过滤条件。

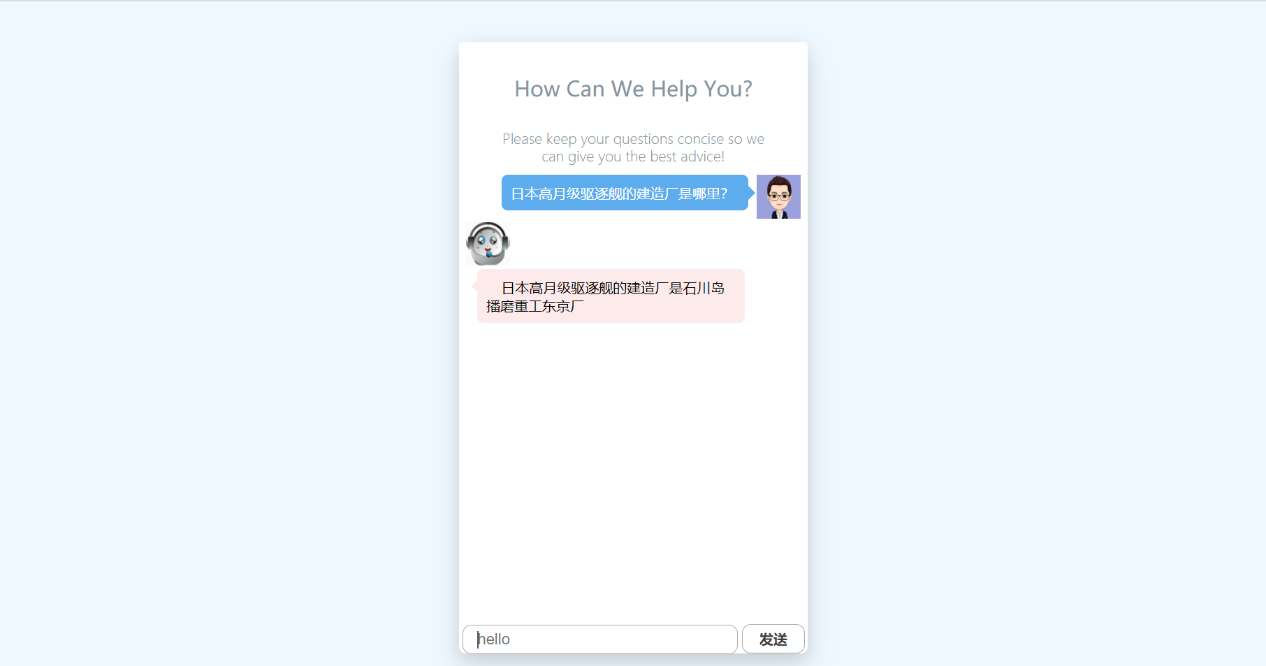
RETURN：返回所需要的。

**3.4 KGQA实践**

本文采用 Python 的 Django Web 开发框架搭建军事武器问答系统平台， 用户通过前端的交互页面获取图谱中的知识， 前端实现主要运用了 html、 javascript 和 CSS （层叠样式表） 等技术。前端交互模块有知识的查询和问答 2部分。

1）知识查询： 用户在前端输入问句， 提交后自动向服务器发送请求 （Requestion） ， 控制层 （Controller） 将请求消息传递给后台服务， 经过后台的问句解析， 解析结果传递到数据层， 由数据层将查询结果返回给用户。

2）知识问答： 类似聊天机器人， 用户输入问题的答案会直接在界面上显示； 如未查到答案， 则返回“对不起， 没有查到结果！ ” 。

****  
  问答系统演示截图