ppt 1

大家好，今天我们汇报的主题是基于知识图谱的问答系统，我们小组成员有黄烁佳和欧阳尚铭

ppt 2

我们的汇报分为两部分内容—一是对项目的介绍 二是项目的不足和改进

ppt3

现在讲解第一部分，项目工作的介绍。

ppt4

项目工作，项目工作分为两部分——金融知识图谱的构建和问答系统的构建

ppt5

一 金融知识图谱的构建

ppt6

对于构建金融知识图谱，首先我们需要从网页中获取所需要的信息，所以我们选择金融网站同花顺F10作为数据源，该网站记录了上市公司的各种信息，如图所示：

ppt7

通过网络爬虫，我们得到了4个结点csv文件和3个关系csv文件，分别为高管、企业、概念、行业结点文件和董事会成员、行业属于、概念属于关系文件，如图所示。

ppt8

得到可以导入neo4j的csv文件后，下面通过neo4j的导入命令将csv文件转换为知识图谱

ppt9

最终得到的金融知识图谱如图所示

ppt10

1. 问答系统的建立。该系统的结构图如图所示:

(这一部分一定要用鼠标来描述)

首先用户的问题通过实体抽取得到{实体类型：实体}字典和实体抽取后的问题，实体抽取后的问题通过属性映射得到问题模板编号，然后里利用之前的得到的字典和问题模板编号来生成相应的查询语句，最后将查询结果进行语言润色得到答案。下面来详细说明每个步骤：

ppt11

用户问题清洗：念ppt就行

ppt12

实体抽取:念ppt就行

ppt13

1. 属性映射:念ppt就行

ppt14

念ppt就行

ppt15：

最后我们用一张实例图来总结以上的流程

ppt16：

项目效果如图所是，比如我输入问题“胡晓东是谁”，会得到”胡晓东，性别女，年龄55岁，先就职于株治集团企业担任独立董事”的答案

ppt17：

接着讲讲第二部分，项目的不足和改进，本来这一部分需要去实现的，但是时间关系，就只是提出想法

ppt18：

念ppt就行了

ppt19：

我们的汇报到此结束，谢谢大家。