
化妆品 知识图谱

小组:5KG

小组成员及分工

胡川

- 知识获取、知识融合、问答系统、可视化、编程实现、报告撰写

陈楚珂

- 知识获取、知识融合、报告撰写

张祥祥

- 知识获取、报告撰写

黄晓辉

- 知识获取、知识融合

张欢

- 知识获取、报告撰写

任务定义

④ 主题

- 化妆品，以口红为主

④ 背景

- 当今社会商品类型繁多，信息分散且价值密度低，如何准确快速地 搜寻满足自己需求的产品成为了日常难题。小组旨在通过建立化妆品的知识图谱，全面且精准地建立相关知识库；基于知识图谱设计智能 问答系统，帮助科学选择符合需求的产品。

④ 任务

- 化妆品知识图谱构建
- 图谱可视化
- 基于知识图谱的问答系统

知识获取



✓ 美妆APP

- 较规范的化妆品数据
- 品牌
- 名称
- 图片
- 颜色
- ...

✓ 电商

- 作为补充数据
- 商品评价
- 商品价格

知识融合和处理

☑ 多源数据的融合

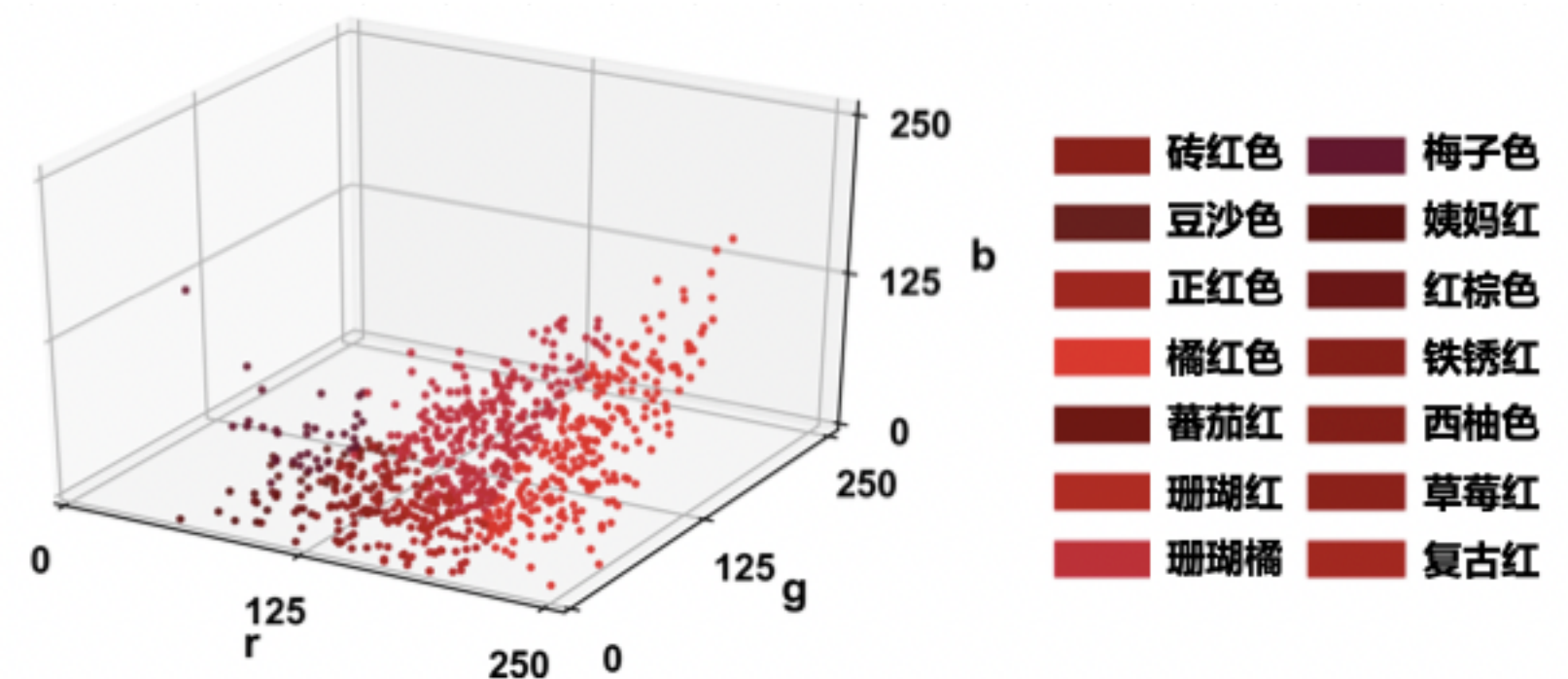
- 根据化妆品的品牌、系列、编号（色号）来对齐实体。

<Dior,28,319,颜色漂亮...> <Dior,28,rgb(164,23,29)...>

<Dior,28,rgb(164,23,29),319,颜色漂亮...>

☑ 颜色聚类

- 使用K-means算法对口红的颜色做聚类



知识存储

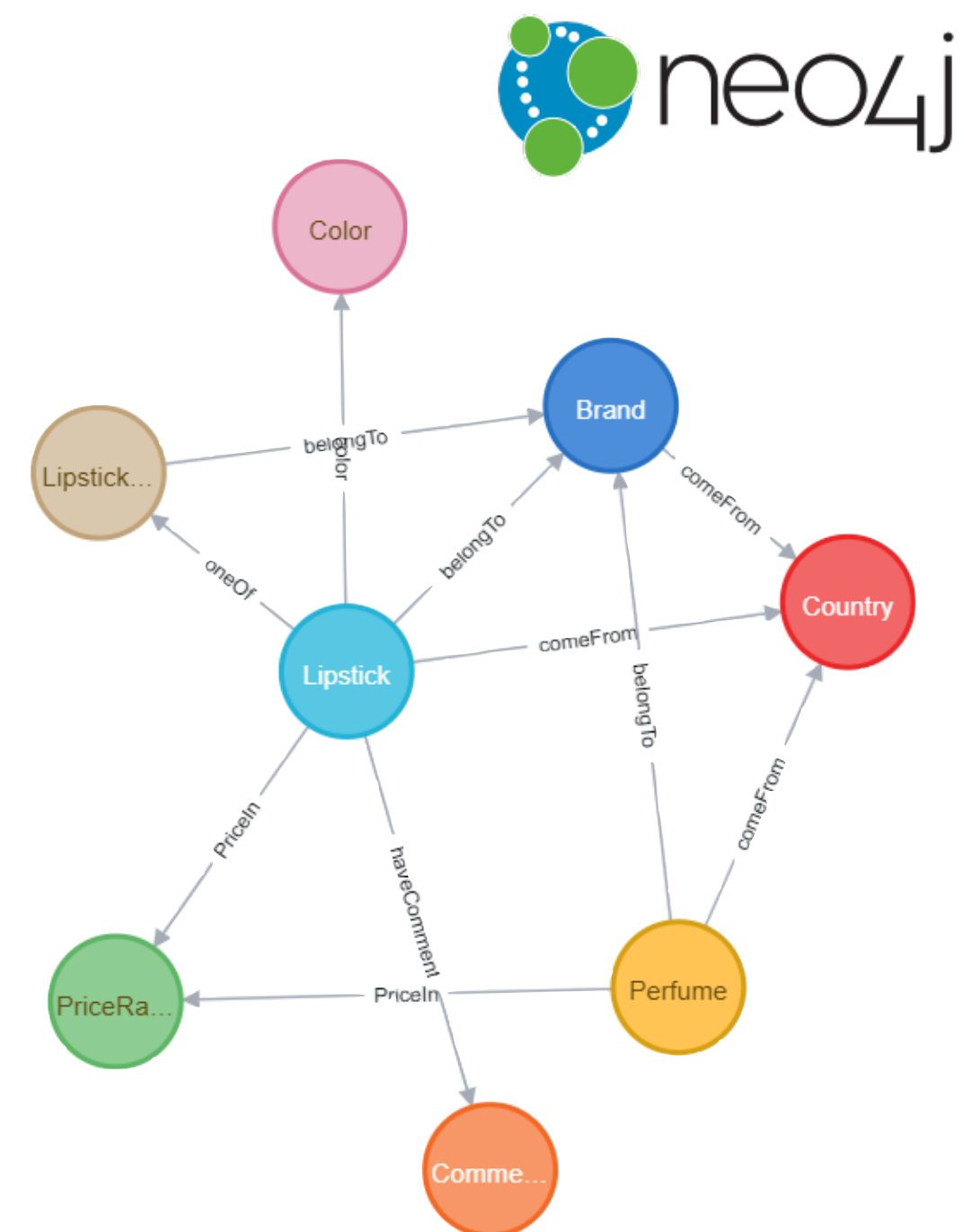
☑ 使用Neo4j存储知识图谱

☑ 实体类型

- 国家
- 品牌
- 香水
- 口红
- 口红系列
- 评论词
- 价格区间
- 色系

☑ 关系

- belongTo
- oneOf
- comeFrom
- priceIn
- haveComment
- color



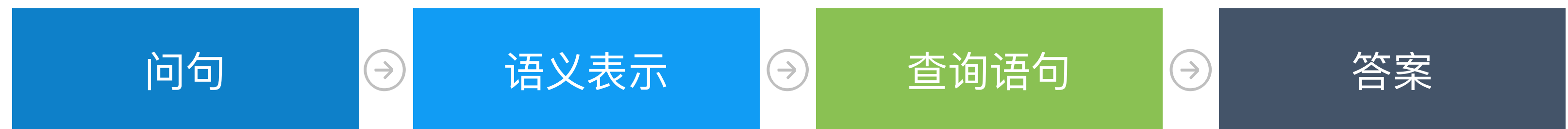
问答系统

☑ 特点

- 结构简单
- 数据量小
- 缺少<问题,答案>数据集

☑ 方法

- 基于符号逻辑的问答方法
- 流水线式



问答系统

✓ 实体链接

- 使用字典匹配完成短句到**实体**的映射

✓ 句法分析

- 使用HanLP中文自然语言处理工具包

✓ 关键词识别

- 识别句法树中的**标签词**、**疑问词**、**属性词**、**属性值**等关键词
- 删除未识别的叶子节点

✓ 问题图转换

- 根据规则删除剩余未识别节点，并把句法树转换为无向图

问句



实体链接



句法分析



关键词识别

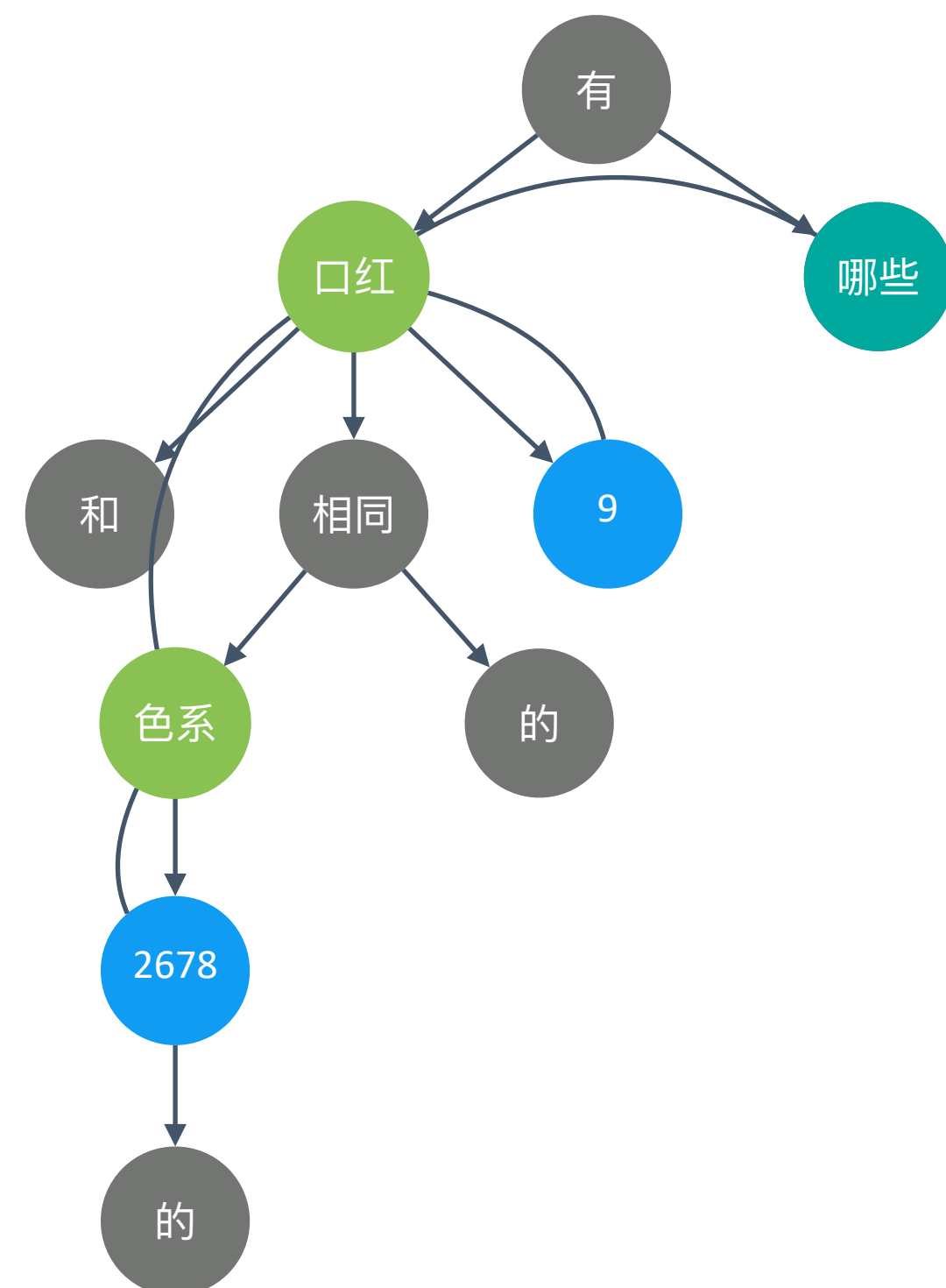


问题图转换

和小羊皮唇膏307的色系相同的迪奥口红有哪些？

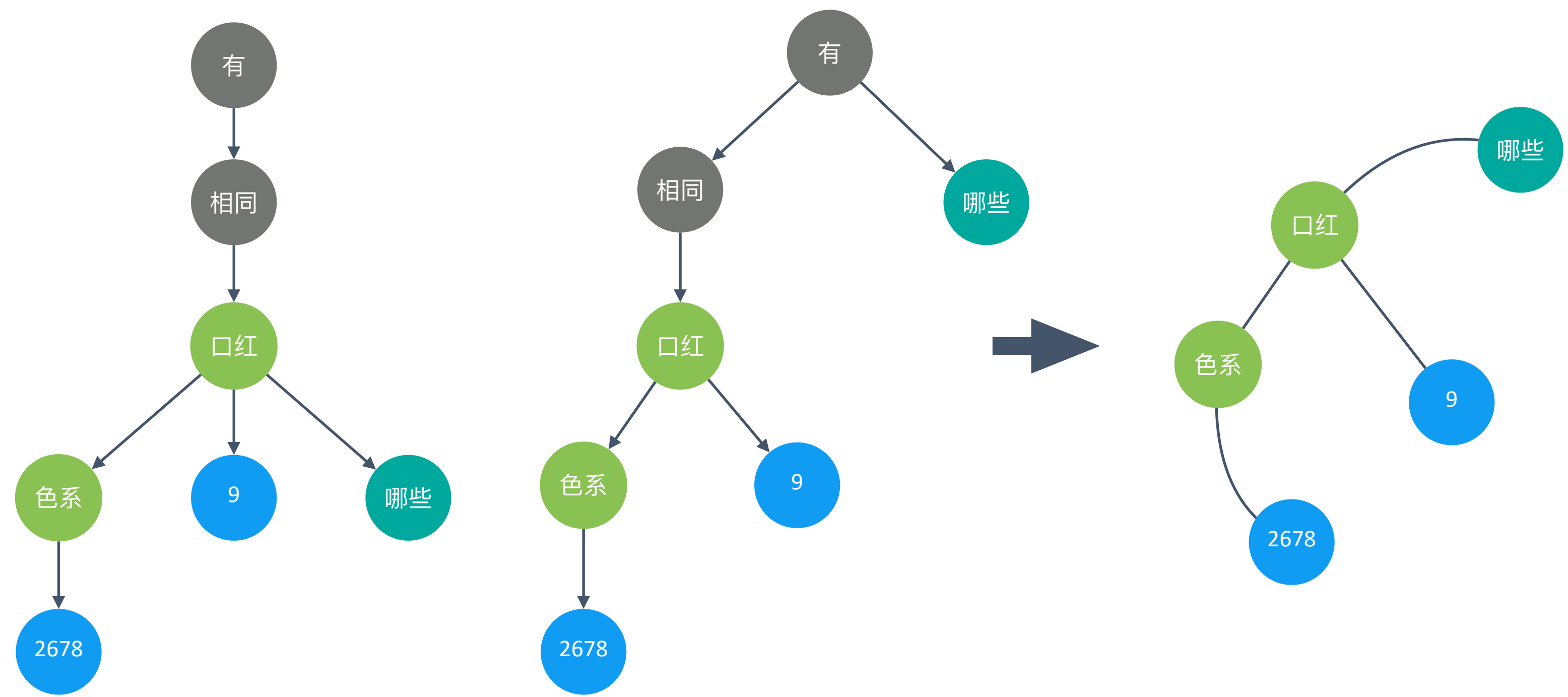


和**小羊皮唇膏307**的色系相同的**迪奥**口红有哪些？
<label=Lipstick, id=2678> <label=Brand, id=9>



问答系统

有哪些迪奥口红和小羊皮唇膏 307 的色系相同?



迪奥口红和小羊皮唇膏 307 的色系相同的有哪些

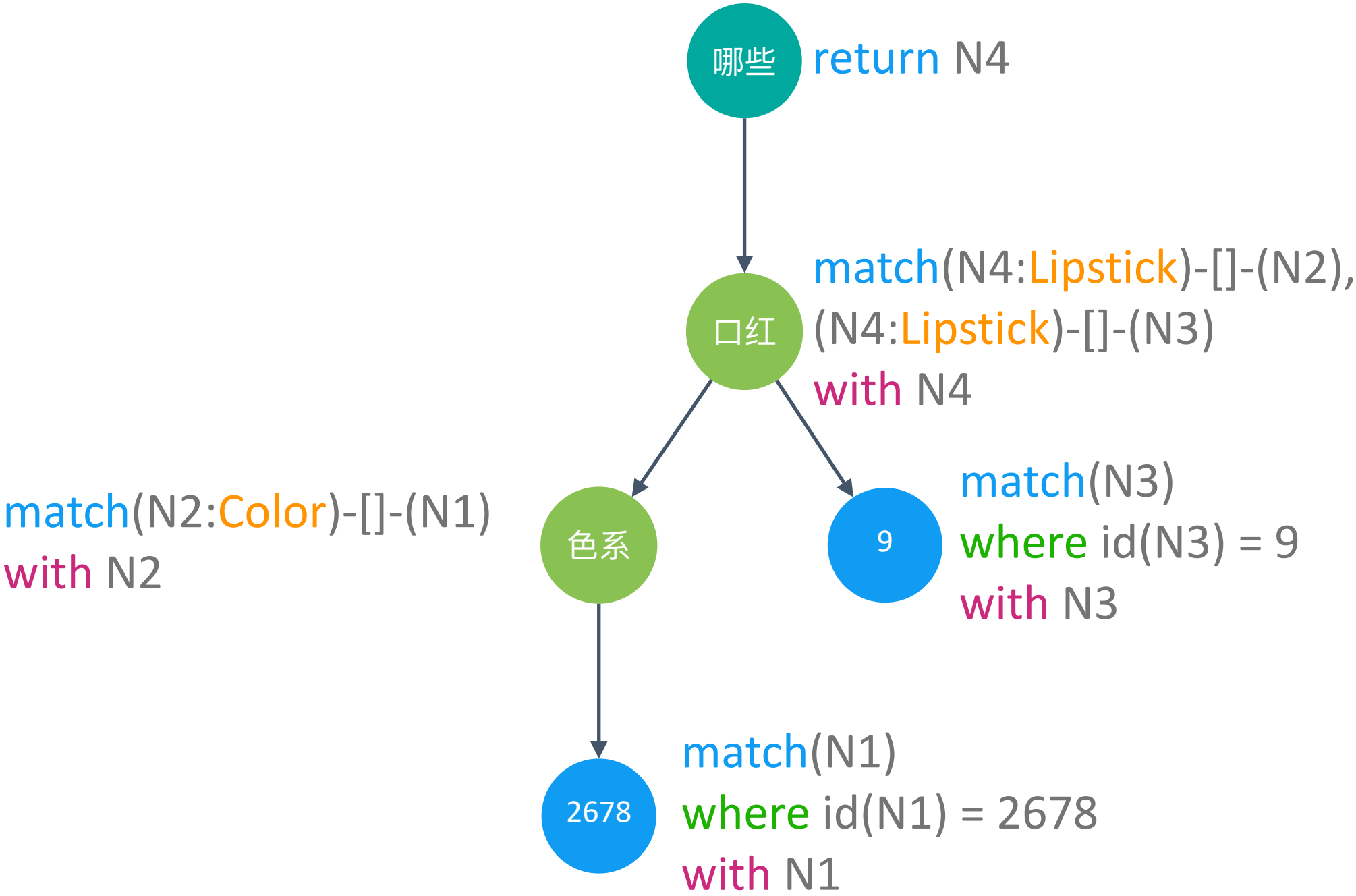
问答系统

问题图解析

- 以问题词为根节点转换成树
- 自底向上
- 解析成Cypher查询语言

查询

- 使用Cypher查询图数据库，返回答案



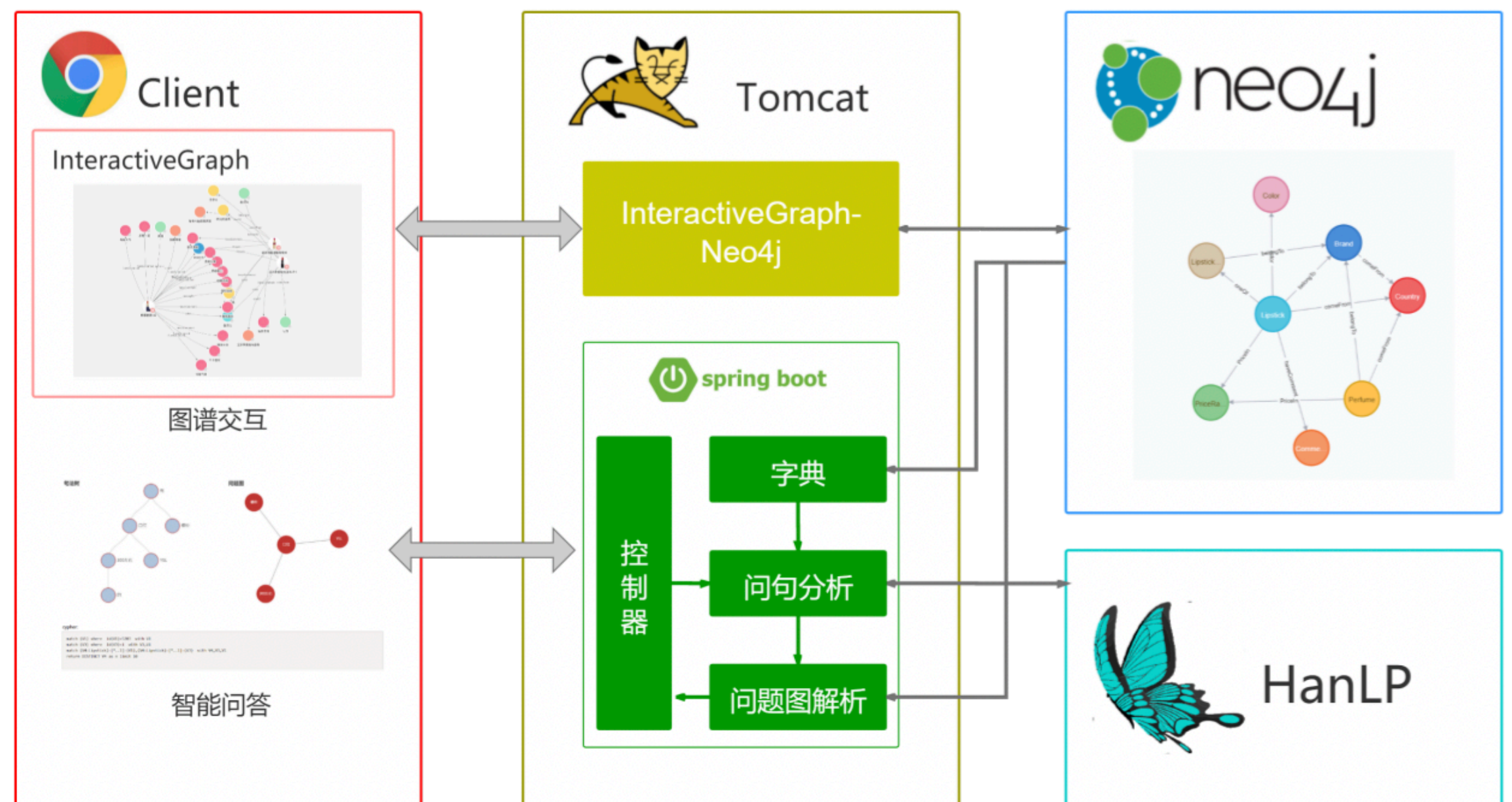
可视化 and 实现

✔ 图谱可视化

- InteractiveGraph(<https://github.com/grapheco/InteractiveGraph>)

✔ 问答系统

- 使用Java实现
- 使用HanLP自然语言处理工具包
- 基于SpringBoot框架实现接口



下一步工作

☑ 短语映射

- 实体链接
 - 关键词识别
 - 对值的识别
- } 完全匹配 => 相似度计算

☑ 问题图生成

- 使用依存关系
- 歧义问题
- 逻辑词丢失（如“相同”）
- 逻辑计算（如“做多”）

演示

☑ 图谱可视化

- <http://60.205.227.226:8080/CosmeticsKG-Web/>

☑ 问答系统

- <http://60.205.227.226:8080/CosmeticsKG-Web/qa.html>



THANKS!

