

# 宝安搜索数据问题

---

## 宝安搜索数据问题

1. 数据重复
2. 数据缺失
3. 地址数据不规范
  - 3.1 去除不合理数据
  - 3.2 数据简化
4. 建议

## 1. 数据重复

---

- 类型不同的节点，名字重复  
例："新安街道"既是社区名也是街道名
- 数据属性值重复  
"厂房"既是楼栋节点某个属性的一个值也是房屋节点某个属性的一个值
- 数据简化  
用户在搜索的时候并不会输入一些较长的属性值，需要对这些较长的属性值做简化

## 2. 数据缺失

---

例：楼栋节点标签缺失“危楼”

## 3. 地址数据不规范

---

### 3.1 去除不合理数据

- 2站
- 南楼
- 5室

### 3.2 数据简化

- 西海岸花园A1A2A3A4A5A6A7栋1A2A3A4A5A6A7204室
- 西海岸花园B1B2B3B4B5B6B7B9栋1、B2、B3、B4、B5、B6、B7、B9栋B207室

## 4. 建议

---

- 统一双方所有使用的地址数据
- 优先做地址数据规范化处理