**58同城(房源)分布式爬虫**

**核心功能和优势**

1. 支持分布式
2. 支持虚拟化并集中部署
3. 不封号
4. 自动化监控
5. 最新最新房源实现分钟级同步
6. 规模化集群，支持水平扩容

**数据类型**

1. 支持抓取100城房源信息
2. 支持手机号识别
3. 支持自动化监听关键词匹配的内容

# 58同城——爬取步骤

<https://blog.csdn.net/jingwenliu/article/details/82120850>

# 58同城数据库架构设计思路

<https://blog.csdn.net/GV7lZB0y87u7C/article/details/78372765>

基于Scrapy框架设计爬虫程序爬取温州租房信息并分析  
  
<https://www.xianjichina.com/news/details_142871.html>

# Scrapy分布式、去重增量爬虫的开发与设计

<https://www.sohu.com/a/255590351_787107>