

文章推荐系统

爬虫部署

urllib.request+Beautifulsoup

爬取、解析

流程控制

清洗、解析

本地数据库SQLite

文章文本

文章url

用户界面

操作选项

搜索结果显示

文本推荐算法

文本推荐SimHash
算法

Hamming Distance