体系结构需求定义

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **编号** | **体系结构需求** | **优先级** | **约束编号** | **设计约束** | **相关约束** |
| R38 | 爬取其他网站的商品信息及价格 | 10 | C49 | 爬取的数据全面且准确 |  |
|  |  |  | C50 | 爬取信息的来源接口要易于更换 |  |
| R39 | 保证爬取数据的实时性、有效性 | 7 | C51 | 系统能够定期检测信息来源网站的数据是否更新，若更新则对系统内数据信息进行更新 |  |
| R40 | 爬取信息兼容 | 5 | C52 | 实现至少2个网站兼容 |  |
|  |  |  | C53 | 技术方案上要更加开放，例如实现语言兼容、单位兼容 | 促进可扩展性 |
| R41 | 爬取兼容模块化 | 6 | C54 | 模块分开，对未来可能出现的兼容内容具有更好的兼容性 | 促进快速开发  促进可扩展性 |
| R42 | 爬取策略可变更 | 5 | C55 | 可以修改爬虫爬取信息的策略 | 促进可扩展性  阻碍C53 |

场景描述

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | | 内容 | |
| 场景ID | | S1 | |
| 商业目标 | | 能够从其他网站获取商品信息、价格信息 | |
| 相关需求与设计约束 | | R38爬取其他网站的商品信息及价格 | C49爬取的数据全面且准确  C50爬取信息的来源接口要易于更换 |
| 场景内容 | 刺激 | 爬取其他购物网站商品 | |
| 刺激源 | 系统或数据库管理员 | |
| 环境 | 正常使用时 | |
| 制品 | 高效的网页信息爬取模块 | |
| 响应 | 更新和扩展已有商品信息 | |
| 响应的度量 | 能够从各种不同的网站上获取商品信息，且扩展性强 | |

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | | 内容 | |
| 场景ID | | S2 | |
| 商业目标 | | 保证数据库内和网页上显示的数据是最新的 | |
| 相关需求与设计约束 | | R39保证爬取数据的实时性、有效性 | C51系统能够定期检测信息来源网站的数据是否更新，若更新则对系统内数据信息进行更新 |
| 场景内容 | 刺激 | 系统到了更新数据的周期 | |
| 刺激源 | 系统或数据库管理员 | |
| 环境 | 正常使用时 | |
| 制品 | 检测计时器和数据更新比对模块 | |
| 响应 | 对数据不同的部分进行快速更新，保证数据库信息的实时性和有效性 | |
| 响应的度量 | 爬取数据的周期在1天左右，服务器对数据更新所需的时间不超过10分钟，更新时不影响网站正常使用 | |

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | | 内容 | |
| 场景ID | | S3 | |
| 商业目标 | | 可以兼容多个网站，包括英文网站 | |
| 相关需求与设计约束 | | R40爬取信息兼容 | C52 至少兼容两个网站 |
| 场景内容 | 刺激 | 新需兼容网站 | |
| 刺激源 | 加盟商家 | |
| 环境 | 正常使用时 | |
| 制品 | 爬取信息兼容模块 | |
| 响应的度量 | 拓展爬取信息兼容子模块，可以兼容新网站 | |

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | | 内容 | |
| 场景ID | | S4 | |
| 商业目标 | | 兼容各个网站的不一致性（定义匹配格式），例如商品名称、价格单位等 | |
| 相关需求与设计约束 | | R40 爬取信息兼容 | C53 技术方案开放 |
| 场景内容 | 刺激 | 开始一次信息爬取 | |
| 刺激源 | 数据库管理员 | |
| 环境 | 正常使用时 | |
| 制品 | 爬取信息兼容模块 | |
| 响应的度量 | 从爬取源网页进行信息爬取并进行兼容操作 | |

|  |  |  |  |
| --- | --- | --- | --- |
| 项目 | | 内容 | |
| 场景ID | | S5 | |
| 商业目标 | | 兼容新的可能出现的或更改现有的爬虫策略 | |
| 相关需求与设计约束 | | R42 爬取策略可变更 | C55可以修改爬虫爬取信息的策略 |
| 场景内容 | 刺激 | 新增爬虫策略 | |
| 刺激源 | 系统管理员 | |
| 环境 | 正常使用时 | |
| 制品 | 爬取信息兼容模块 | |
| 响应的度量 | 拓展爬取信息兼容模块，可以兼容新爬虫策略 | |

设计决策

设计决策1

|  |  |  |
| --- | --- | --- |
| 项目 | 内容 |  |
| 体系结构需求  和约束 | R38爬取其他网站的商品信息及价格 | C49爬取的数据全面且准确  C50爬取信息的来源接口要易于更换 |
| 方案 | 封装页面爬取算法，然后对多个购物网站进行信息爬取 | |
| 影响 | 逻辑视图、开发视图 | |
| 详细设计约束 | 无 | |

设计决策2

|  |  |  |
| --- | --- | --- |
| 项目 | 内容 |  |
| 体系结构需求  和约束 | R39保证爬取数据的实时性、有效性 | C51系统能够定期检测信息来源网站的数据是否更新，若更新则对系统内数据信息进行更新 |
| 方案 | 将爬取的数据与数据库中的数据进行比对，对新数据进行暂存，然后集中更新；  更新时选择在流量少的时段 | |
| 影响 | 逻辑视图、开发视图 | |
| 详细设计约束 | 无 | |

设计决策3

|  |  |  |
| --- | --- | --- |
| 项目 | 内容 |  |
| 体系结构需求  和约束 | R40 爬取信息兼容 | C52 至少实现2个网站兼容  C53 技术方案上要更加开放 |
| 方案 | 封装每个网站的兼容技术方案 | |
| 影响 | 逻辑视图、开发视图 | |
| 详细设计约束 | 隐藏兼容性实现细节 | |

设计决策4

|  |  |  |
| --- | --- | --- |
| 项目 | 内容 |  |
| 体系结构需求  和约束 | R41 爬取兼容模块化 | C54 模块分开，应对新增兼容内容 |
| 方案 | 分模块实现对于语言、价格、单位的兼容 | |
| 影响 | 逻辑视图、开发视图 | |
| 详细设计约束 | 隐藏兼容性实现细节 | |

设计决策5

|  |  |  |
| --- | --- | --- |
| 项目 | 内容 |  |
| 体系结构需求  和约束 | R42 爬取策略可变更 | C53 技术方案上要更加开放  C55 修改爬虫爬取信息的策略 |
| 方案 | 使用strategy模式 | |
| 影响 | 逻辑视图、开发视图 | |
| 详细设计约束 | 隐藏爬取策略实现细节 | |