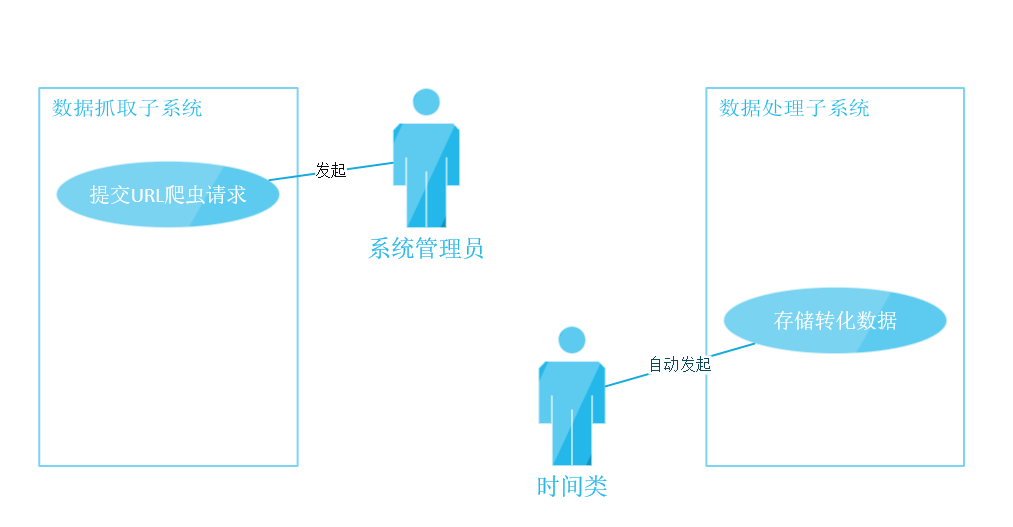
用例清单

|  |  |  |
| --- | --- | --- |
| 用例名称 | 用例描述 | 参与者 |
| 提交URL爬虫请求 | 系统管理员在系统中输入要爬取数据的页面的URL，获取爬虫原始数据。 | 系统管理员 |
| 存储转化数据 | 数据处理子系统将爬虫原始数据整理并写入数据库中储存。 | 时间类（Time） |

用例图



用例的详细描述

**数据抓取子系统**

**Author (s): \_\_\_\_\_\_\_Wen Liang\_\_\_\_\_\_\_\_ Date: \_\_2016/09/26\_\_\_\_\_\_\_\_**

**Version: \_\_\_\_\_1.0\_\_\_\_\_\_**

|  |  |  |
| --- | --- | --- |
| **USE CASE NAME:** | 提交URL爬虫请求 | **USE CASE TYPE** |
| **USE CASE ID:** |  | Abstract: **🞏** |
| **PRIORITY:** | 高 | Extension: **🗹** |
| **INVOKED BY:** | 无 |  |
| **PARTICIPATING ACTORS:** | 系统管理员 | |
| **DESCRIPTION:** | 系统管理员在系统中输入要爬取数据的页面的URL，获取爬虫原始数据。 | |
| **PRE-CONDITION:** | 对于该网页的爬虫逻辑及代码已初始化完成 | |
| **TYPICAL COURSE**  **OF EVENTS:** | **Step 1**: 系统管理员进入【数据抓取】界面，输入将要爬取的网页的URL | |
|  | **Step 2**: 点击【检查】按钮检查系统中是否有对应URL的爬虫脚本，并返回对应的提示对话框（验证通过或重新输入）。 | |
|  | **Step 3**: 系统中有对应URL的爬虫脚本，验证通过，点击【确定】返回上一级菜单后点击【开始爬取】，系统开始爬取网页，返回原始数据 | |
| **ALTERNATE COURSES:** | **Step 3 :** 系统中没有对应的爬虫脚本，点击【确定】重新输入URL或退出系统 | |
| **POST-CONDITION:** | 对爬取的数据确认保存、丢弃或重新爬取 | |

**数据处理子系统**

**Author (s): \_\_\_\_\_\_\_Wen Liang\_\_\_\_\_\_\_\_ Date: \_\_2016/09/26\_\_\_\_\_\_\_\_**

**Version: \_\_\_\_\_1.0\_\_\_\_\_\_**

|  |  |  |
| --- | --- | --- |
| **USE CASE NAME:** | 存储转化数据 | **USE CASE TYPE** |
| **USE CASE ID:** |  | Abstract: **🞏** |
| **PRIORITY:** | 高 | Extension: **🗹** |
| **INVOKED BY:** | 无 |  |
| **PARTICIPATING ACTORS:** | 时间类 | |
| **DESCRIPTION:** | 数据处理子系统将爬虫原始数据整理并写入数据库中储存。 | |
| **PRE-CONDITION:** | 系统管理员对于数据爬取子系统的原始数据已确认保存 | |
| **TYPICAL COURSE**  **OF EVENTS:** | **Step 1**: 系统按照预设的正则表达式对原始数据进行匹配和筛选，生成原始表单 | |
|  | **Step 2**: 系统对原始表单进行整理（生成ID、拆分成若干对应的表等） | |
|  | **Step 3**: 系统将整理后的表单保存到数据库中，返回【数据处理完毕】提示对话框 | |
| **ALTERNATE COURSES:** | **无** | |
| **POST-CONDITION:** | 对处理后的数据，选取相关的分析规则进行查询和分析。 | |

非功能性需求

1.效率： 数据爬取时间与写入数据库时间不能过长