# 10附录

## 数据字典

|  |  |
| --- | --- |
| 数据结构卡片编号：1  数据结构名称：网页元素  别名：原新闻 | |
| 说明：爬虫程序从网页直接爬取到的数据。 | 有关数据流/数据储存：  从网页上爬取下来  信息容量：10000条 |

|  |  |
| --- | --- |
| 数据结构卡片编号：2  数据结构名称：新闻分类表  别名：不同的新闻 | |
| 说明：每条新闻都带有相关的分类信息。分类标签 = 标签ID+标签名字+其他相关标签。 | 有关数据流/数据储存：  存储到Mongo DB中的新闻表中  信息容量：10000条 |

|  |  |
| --- | --- |
| 数据结构卡片编号：3  数据结构名称：喜欢的新闻  别名：我的喜欢表 | |
| 说明：这是用户点击“喜欢”按钮时产生的数据。  喜欢信息 = 喜欢的时间 + 但用户id + 新闻的id + 新闻的标签 | 有关数据流/数据储存：  从用户接口到「我的喜欢表」  信息容量：5000条 |

|  |  |
| --- | --- |
| 数据结构卡片编号：4  数据结构名称：历史记录表  别名： | |
| 说明：这是用户每次浏览新闻时产生的数据  历史记录 = 用户id + 阅读时间 + 新闻id + 新闻标签 | 有关数据流/数据储存：  从用户接口到「我的阅读历史表」  信息容量：20000条 |