数据库选型：MySQL 8.0.40

**用户表t\_user**支撑【注册登录】功能

**作用：** 存储用户的登录注册信息。

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名 (Column)** | **数据类型 (Type)** | **长度/约束 (Constraints)** | **描述 (Description)** |
| **id** | BIGINT | PRIMARY KEY, AUTO\_INCREMENT | 用户唯一ID，主键，自增 |
| username | VARCHAR | 255, NOT NULL, UNIQUE | 用户名，必须唯一，不能为空 |
| password | VARCHAR | 255, NOT NULL | 加密后的密码，不能为空 |
| create\_time | DATETIME | NOT NULL | 账户创建时间 |
| update\_time | DATETIME | NOT NULL | 账户信息最后更新时间 |

SQL语句

CREATE TABLE `t\_user` (

`id` BIGINT NOT NULL AUTO\_INCREMENT COMMENT '用户ID',

`username` VARCHAR(255) NOT NULL COMMENT '用户名',

`password` VARCHAR(255) NOT NULL COMMENT '加密后的密码',

`create\_time` DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP COMMENT '创建时间',

`update\_time` DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP ON UPDATE CURRENT\_TIMESTAMP COMMENT '更新时间',

PRIMARY KEY (`id`),

UNIQUE INDEX `uk\_username` (`username`)

) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COMMENT='用户表';

**爬取历史表 t\_crawl\_history**支撑【爬取普通历史展示】功能

**作用：** 记录每个用户的爬取操作历史。

**t\_crawl\_history (扩展后)**

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名** | **数据类型** | **约束** | **描述** |
| **id** | BIGINT | PK, AI | 历史记录ID |
| user\_id | BIGINT | NOT NULL | 用户ID |
| **crawl\_type** | VARCHAR(20) | NOT NULL | 爬取类型。例如：'SINGLE\_URL', 'KEYWORD\_BATCH' |
| url | TEXT |  | 如果是单URL爬取，这里记录URL。批量爬取时可为空。 |
| title | VARCHAR(255) |  | 如果是单URL爬取，记录标题。批量爬取时，可以记录类似“关键词[xxx]的爬取任务” |
| **params** | JSON |  | 用于存储复杂参数。如果是批量爬取，这里存{"keyword": "xxx", "startDate": "...", "endDate": "..."} |
| crawl\_time | DATETIME | NOT NULL | 爬取操作的时间 |

CREATE TABLE `t\_crawl\_history` (

`id` BIGINT NOT NULL AUTO\_INCREMENT COMMENT '历史记录ID',

`user\_id` BIGINT NOT NULL COMMENT '用户ID',

`crawl\_type` VARCHAR(20) NOT NULL COMMENT '爬取类型: SINGLE\_URL, KEYWORD\_BATCH',

`url` TEXT NULL COMMENT '单URL爬取时的目标URL',

`title` VARCHAR(255) NULL COMMENT '任务标题或描述',

`params` JSON NULL COMMENT '复杂查询参数（如关键词、日期）',

`crawl\_time` DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP COMMENT '爬取时间',

PRIMARY KEY (`id`),

INDEX `idx\_user\_id` (`user\_id`)

) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COMMENT='爬取历史表';

**新闻数据表 t\_news\_data** 支撑【新闻数据统计分析】功能

**作用：** 用于存储“定时任务”或“按关键词批量爬取”所获取的新闻数据。

|  |  |  |  |
| --- | --- | --- | --- |
| **字段名 (Column)** | **数据类型 (Type)** | **长度/约束 (Constraints)** | **描述 (Description)** |
| **id** | BIGINT | PRIMARY KEY, AUTO\_INCREMENT | 新闻唯一ID，主键，自增 |
| url | VARCHAR | 768, NOT NULL, UNIQUE | 新闻的原始URL，必须唯一，用于去重 |
| title | VARCHAR | 255 | 新闻标题 |
| source | VARCHAR | 100 | 新闻来源 (如：新浪新闻) |
| publish\_time | DATETIME |  | 新闻发布时间 |
| content | LONGTEXT |  | 新闻正文内容（HTML或纯文本） |
| keywords | VARCHAR | 255 | 新闻关键词，用逗号分隔 |
| fetch\_time | DATETIME | NOT NULL | 数据抓取入库的时间 |

CREATE TABLE `t\_news\_data` (

`id` BIGINT NOT NULL AUTO\_INCREMENT COMMENT '新闻ID',

`url` VARCHAR(768) NOT NULL COMMENT '新闻原始URL',

`title` VARCHAR(255) NULL COMMENT '新闻标题',

`source` VARCHAR(100) NULL COMMENT '新闻来源',

`publish\_time` DATETIME NULL COMMENT '发布时间',

`content` LONGTEXT NULL COMMENT '新闻正文',

`keywords` VARCHAR(255) NULL COMMENT '关键词',

`fetch\_time` DATETIME NOT NULL DEFAULT CURRENT\_TIMESTAMP COMMENT '抓取时间',

PRIMARY KEY (`id`),

UNIQUE INDEX `uk\_url` (`url`)

) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COMMENT='新闻数据表';