

外部API接口规范

1. 刘老师爬虫API

接口地址

POST http://liu-teacher-api/crawl-data

请求参数

```
{
  "task_id": 123,
  "sources_config": {
    "presetSources": [
      {
        "id": "wechat-group-001",
        "name": "AI技术讨论群",
        "category": "wechat",
        "type": "微信|企微"
      },
      {
        "id": "official-account-002",
        "name": "科技前沿公众号",
        "category": "wechatOfficial",
        "type": "公众号"
      }
    ],
    "customSources": [
      {
        "id": "custom-web-001",
        "name": "自定义网站",
        "url": "https://example.com",
        "category": "web",
        "type": "网页"
      }
    ]
  }
}
```

```
    "webSearchEnabled": true
  },
  "current_execution_time": "2025-09-19T17:42:08.123456+08:00",
  "last_execution_time": "2025-09-18T17:42:08.123456+08:00",
  "web_search_enabled": true,
  "requirement": "用户输入的需求描述"
}
```

响应格式

```
{
  "code": 200,
  "message": "爬取成功",
  "data": {
    "crawler_task_id": "crawler-task-123-1726756928",
    "crawled_data": [
      {
        "source_id": "wechat-group-001",
        "source_name": "AI技术讨论群",
        "source_type": "wechat",
        "content": "爬取到的完整文本内容...",
        "metadata": {
          "crawl_time": "2025-09-19T17:42:08+08:00",
          "data_count": 150,
          "time_range": {
            "start": "2025-09-18T17:42:08+08:00",
            "end": "2025-09-19T17:42:08+08:00"
          }
        }
      }
    ]
  }
}
```

接口说明

- **目的:** 根据时间戳进行增量数据爬取，确保无遗漏
- **时间戳:** `current_execution_time` - `last_execution_time` 之间的所有数据
- **返回数据:** 爬取到的完整长文本内容，供报告生成使用

2. 王璞报告生成API

接口地址

POST <http://wangpu-api/generate-report>

请求参数

```
{
  "task_id": 123,
  "requirement": "用户输入的需求描述",
  "crawler_data": [
    {
      "source_id": "wechat-group-001",
      "source_name": "AI技术讨论群",
      "source_type": "wechat",
      "content": "完整的爬虫数据内容..."
    }
  ],
  "web_search_enabled": true
}
```

响应格式

```
{
  "code": 200,
  "message": "报告生成任务已创建",
  "data": {
    "report_task_id": "report-task-123-1726756928",
    "status": "processing"
  }
}
```

接口说明

- 目的: 根据爬虫数据和用户需求生成智能报告
- 输入: 刘老师返回的爬虫数据 + 用户需求
- 输出: 报告生成任务ID, 后续通过回调获取结果

3. 报告完成回调API (我方提供给王璞)

接口地址

```
POST http://our-api/report-callback
```

请求参数 (王璞调用)

```
{
  "report_task_id": "report-task-123-1726756928",
  "task_id": 123,
  "status": "completed",
  "report_data": {
    "title": "AI技术发展趋势报告",
    "content": "完整的报告内容...",
    "summary": "报告摘要",
    "generated_time": "2025-09-19T17:45:08+08:00"
  }
}
```

响应格式 (我方返回)

```
{
  "code": 200,
  "message": "报告接收成功"
}
```

4. 实现状态

当前状态 (模拟)

-  接口规范已定义
-  时间戳逻辑已实现
-  数据格式已规范化
-  使用模拟数据推进开发

待对接

-  刘老师爬虫API实现
-  王璞报告生成API实现
-  报告回调接收实现

下一步

1. 完善定时任务调度系统
2. 实现报告回调接收接口
3. 优化错误处理和重试机制
4. 准备接口对接测试