# **北京信息科技大学课程实践报告（2022）**

报告题目：开源项目建设发布

课程名称：开源软件开发

任课教师：李宁

项目导师：李宁

小组组长：王鹏

## **团队成员**

王鹏-2021011268-大数据2101

王帅-2021011265-大数据2101

秦昊然-2021011264-大数据2101

高凌屹-2021011263-大数据2101

张宇鑫-2021011255-大数据2101

## **报告要求**

建立一个开源软件项目，并在项目托管平台Github上发布。

报告内容包括以下方面:

\*项目名称

\*项目描述

\*项目许可证

\*项目人员和分工

\*项目开发流程

\*代码库，网站，Wiki，邮件列表，缺陷追踪系统，文档编制

\*项目计划（版本、用户、推广)

\*其他

报告篇幅为1000字以上。格式清晰完整，标点正确，无错别字。

## **项目计划**

以下各项均为必填

项目名称

Excel的自然语言访问接口设计与实现

项目描述

设计与实现字处理文档的自然语言访问接口(段落、标题、图题、表题，表格)

（1）在文档检索等应用中，我们往往需要通过人机友好的方式获取文档中的各排版元素的内容。 例如：“请告诉我某个电子表格文件中有几个sheet??”,“某个电子表格文件中第2个 sheet 的名字是什么？”，“某个电子表格文件中有没有图表？”

（2）在不使用Office软件打开文字处理文档的情况下，机器能够通过分析文档回答类似上述的问题。

许可证（以及为什么选择这个许可证）

MIT

因为这个许可证的授权较宽松，更利于项目的传播与发展。

人员（仅需姓名）和分工

王鹏-项目文档和项目发布+代码开发（sheet数据统计，表格数据读取等工作）

王帅-代码开发（判断表格中是否存在图片）

秦昊然-PPT撰写

高凌屹-代码开发（读取表格中的图片）

张宇鑫-代码开发（统计表格中存在多少图片）

开发流程（项目构思，环境部署，代码编写，代码测试，文档撰写，项目发布等过程）

项目构思（1）：

首先我们需要知道实现字处理文档的自然语言访问接口需要用到什么，通过老师那里我们得知需要用到apache poi这一java依赖包，通过apache 官网可知apache poi 是java的一种API 通过poi 这一依赖可以实现在不打开Office 的情况下读取表格。

环境部署：

下载并安装svn或git、Java JDK 1.8+和Apache Ant 1.8+或Gradle

查看svn或git的代码

将项目导入Eclipse或您最喜欢的IDE

项目构思（2）：

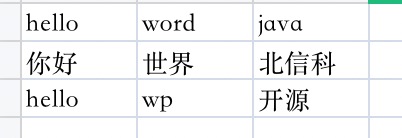
我们将代码的程序分成了四部分，1.分别为sheet的统计与输出，行数的统计与表格数据的输出；2.判断表格中是否存在图片；3.统计该表格中存在了多少图片；4.输出该表格的具体图片。我们采取四部分同时前进的方式，最后进行代码整合

项目构思（3）：

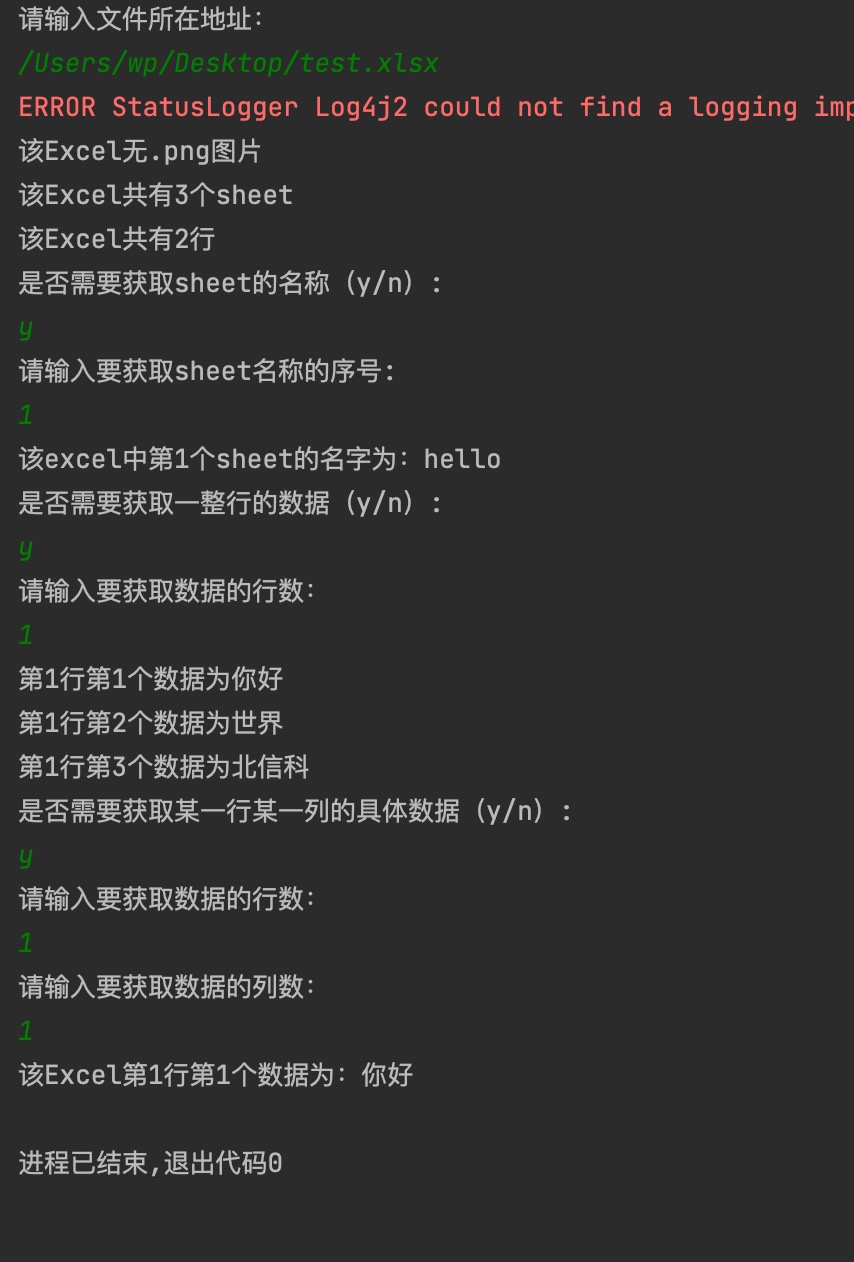
我们通过FileInputStream方法将表格中的数据导入到java中，获得数据流，并通过poi依赖中带的方法对数据流进行修改，以此达到输出特定需求对数据的目的。

代码测试：

测试文件1:

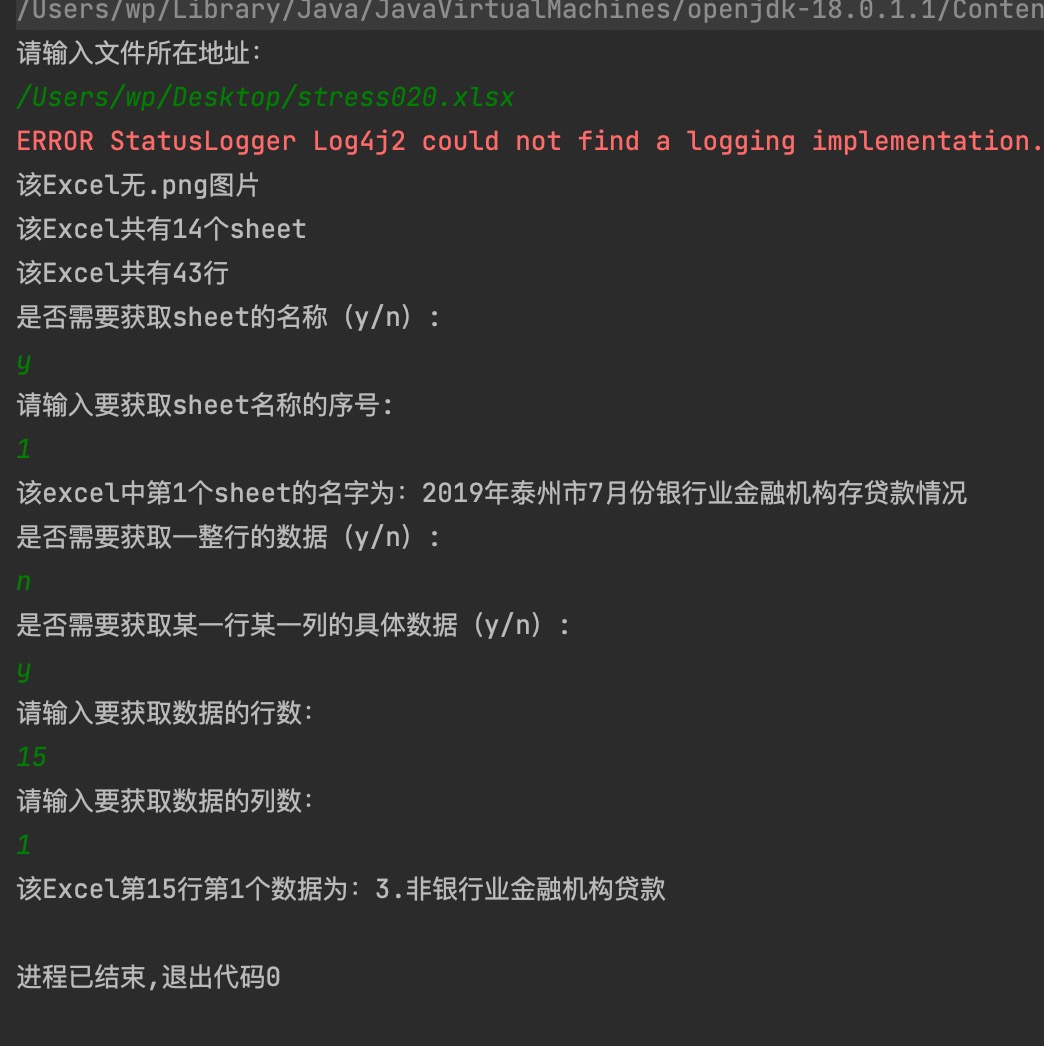


测试结果1:



测试结果2:



测试结果

测试文档3:



测试结果3:



项目信息（给出各项所用的平台、软件及地址）

代码库：<https://github.com/Bistu-OSSDT-2022/9-lin-wangp>

网站：GitHub

Wiki：无

邮件列表：无

缺陷追踪系统：无

文档：https://github.com/Bistu-OSSDT-2022/9-lin-wangp

项目计划（版本发布的计划、面向的用户、如何推广)

以后有时间我们会对程序进一步完善，如使程序可以在读取图片的同时读取表格数据，并且程序的代码没有进行统一，还没有进行界面的程序编写等，这些都需要后续进行完善。

在代码完善的足够好之前我们并不考虑推广的事情，不过我们可以先寻求老师，同学等代码代码大牛的帮助，帮助我们规范代码，设计界面等事情，在未来可能会通过打广告等形式线上线下进行推广。

## **项目文档**

以下内容无需在此填写，请将内容填写到代码仓库中的对应文件。

如果没有内容则文件可以为空，但必须有这些文件存在。

INSTALL

配置、编译和安装该项目的说明信息

FAQ

关于本项目的纯文本格式的常见问题解答。

CREDITS

本项目所有贡献者的列表。

HISTORY

本项目的历史发展演变记录。

READMELICENSE

本项目的许可证条款文件。

本项目的说明。

## **项目开发**

选择一个熟悉的语言开发项目。参考课堂派上提供的范例。

遵守课程讲义中的注意点：

* 按规则命名事物
* 注重编码风格
* 选好工具、语言和平台
* 写好文档
* 做好发布前的准备

本部分无需填写。代码直接提交到代码仓库。

## **项目发布**

1. 注册Github账号。
2. 建立Github代码仓库。
3. 将项目代码发布到Github。
4. 发布项目文档到Github。建议使用Markdown格式。
5. 项目成员的协作需要体现到Github上，也就是小组中的每个成员要加到项目里，并都要有提交。
6. 请在以下填写你发布后的github项目地址（必须），项目必须为公开项目。

<https://github.com/>

## **参考文献**

请列出本报告参考的资料来源。

依赖下载地址：

<https://github.com/apache/poi>

部分代码参考：<https://blog.csdn.net/ITarmi/article/details/123773792?spm=1001.2014.3001.5502>

<https://blog.csdn.net/qq_21137441/article/details/79226171?ops_request_misc=%257B%2522request%255Fid%2522%253A%2522165648545716782246488721%2522%252C%2522scm%2522%253A%252220140713.130102334.pc%255Fblog.%2522%257D&request_id=165648545716782246488721&biz_id=0&utm_medium=distribute.pc_search_result.none-task-blog-2~blog~first_rank_ecpm_v1~hot_rank-3-79226171-null-null.nonecase&utm_term=poi&spm=1018.2226.3001.4450>