目录

[1. 背景 3](#_Toc113432367)

[2. 准备 3](#_Toc113432368)

[2.1 Windows系统安装 3](#_Toc113432369)

[2.1.1 安装JDK8 3](#_Toc113432370)

[2.1.2 安装Pandoc 5](#_Toc113432371)

[2.2 Mac系统安装 7](#_Toc113432372)

[2.2.1 安装JDK8 7](#_Toc113432373)

[2.2.2 安装Pandoc 8](#_Toc113432374)

[3. 运行 10](#_Toc113432375)

[3.1 Windows上运行 10](#_Toc113432376)

[3.2 Mac上运行 11](#_Toc113432377)

[4. 文档优化 11](#_Toc113432378)

[目录的优化： 11](#_Toc113432379)

[采用自动生成目录 11](#_Toc113432380)

[表格的优化 12](#_Toc113432381)

[转换表格类型一 12](#_Toc113432382)

[转换表格类型二 13](#_Toc113432383)

[处理复杂表格 14](#_Toc113432384)

[图片的优化 15](#_Toc113432385)

[其他优化 15](#_Toc113432386)

[5. Q&A 15](#_Toc113432387)

# 背景

# 准备

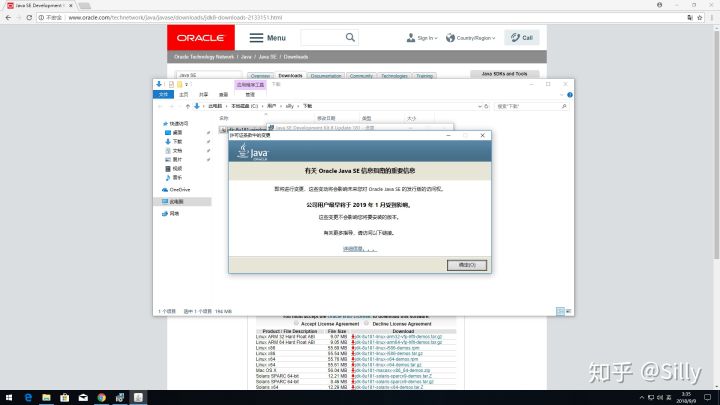
## Windows系统安装

### 安装JDK8

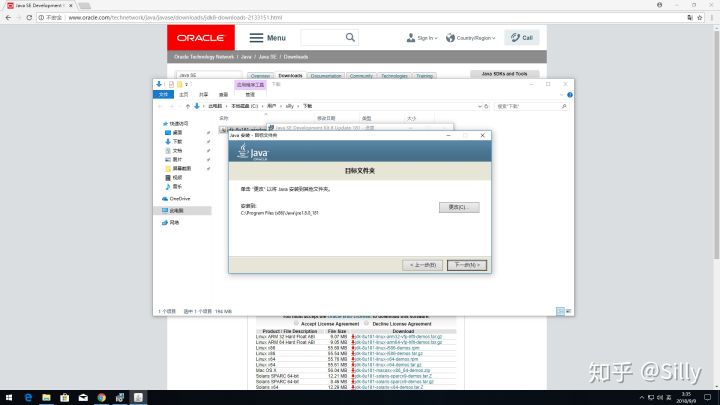
下载地址：

下载jdk-8u341-windows-x64.exe

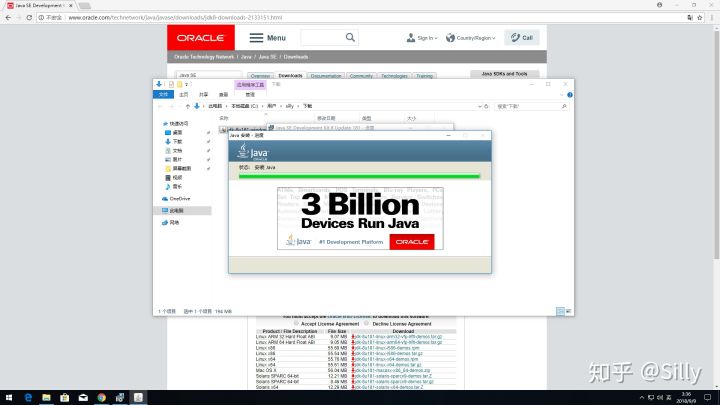
1. 运行jdk-8u341-windows-x64.exe，点击确定



1. 点击更改可以选择安装路径，选择完成后记住安装的路径，点击下一步



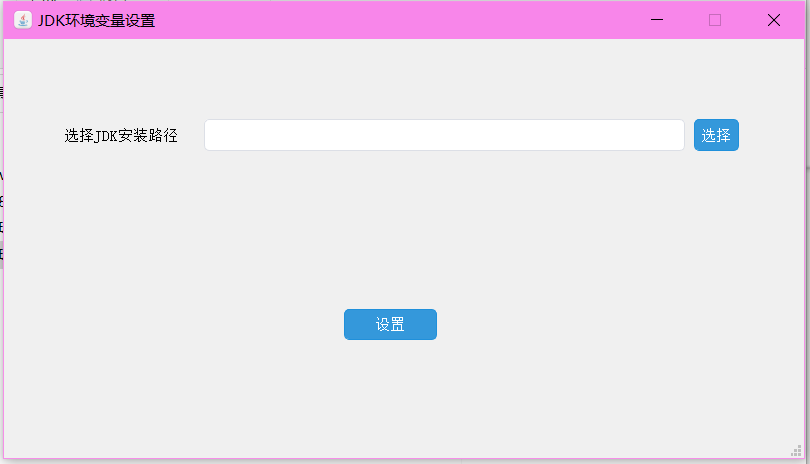
1. 安装中……



1. 安装完成，点击关闭
2. 运行环境变量配置工具（二选一）

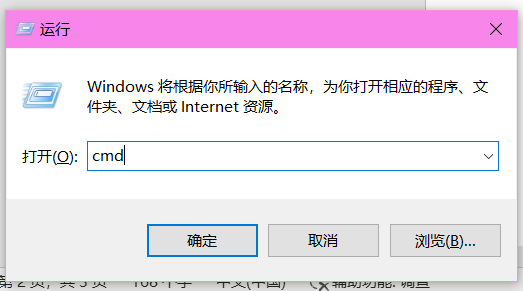


1. 选择找到刚刚安装时的路径，点击设置，设置成功即可



1. 查看jdk是否安装成功

按win + r调出运行窗口，输入cmd调出命令行窗口



在命令行中输入java -version后，如果看到如下，则安装jdk成功



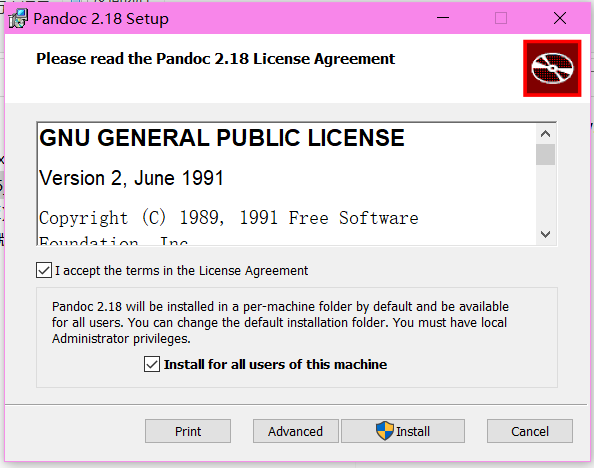
### 安装Pandoc

下载地址：<https://www.pandoc.org/installing.html>

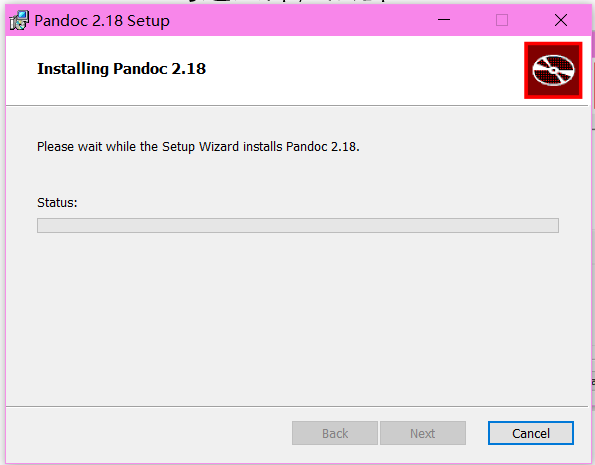
1. 下载后双击安装程序



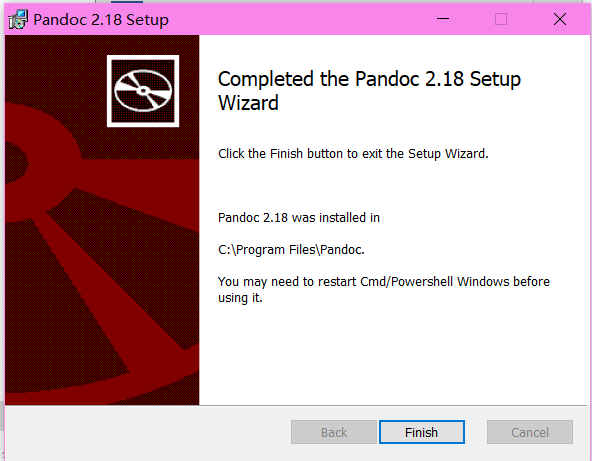
1. 勾选全部，点击Install



1. 等待……



1. 点击Finish完成安装

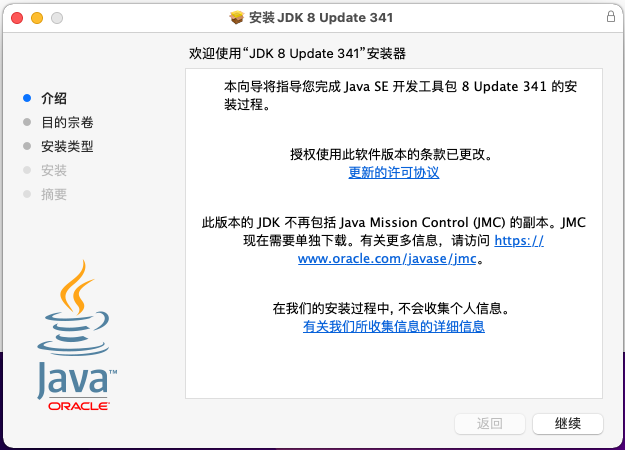


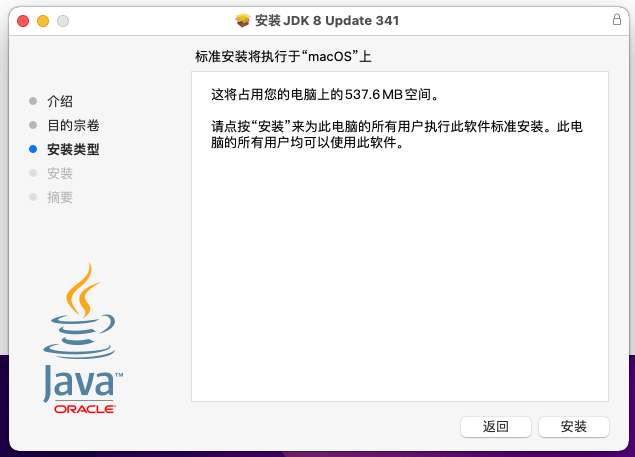
## Mac系统安装

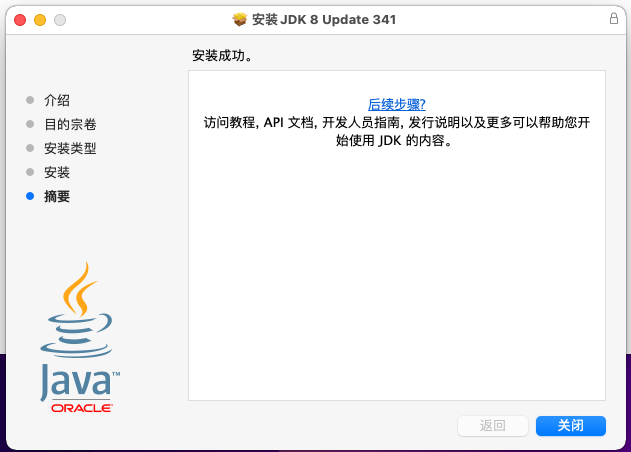
### 安装JDK8

下载地址：

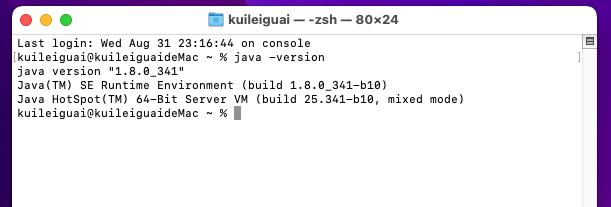
1. 运行下载的安装包,一直点击继续





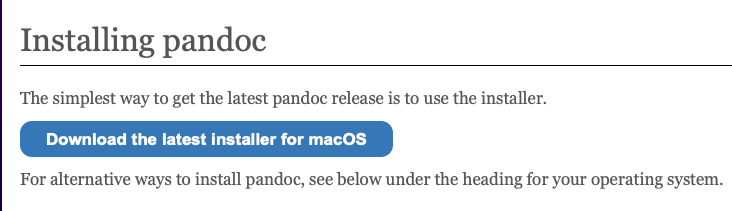


1. 运行终端输入java -version查看是否安装成功



### 安装Pandoc

下载地址:https://www.pandoc.org/installing.html



1. 运行安装包



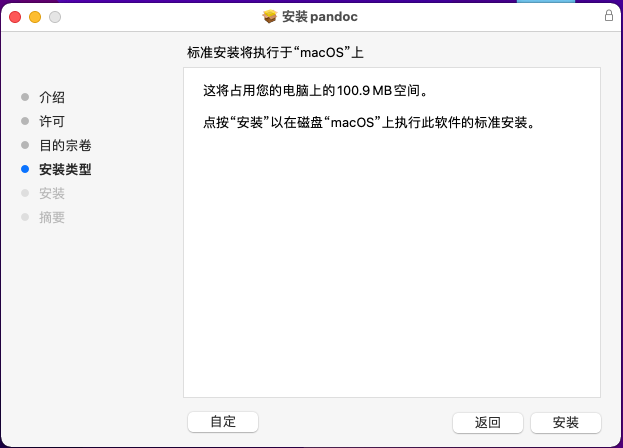
1. 点击继续



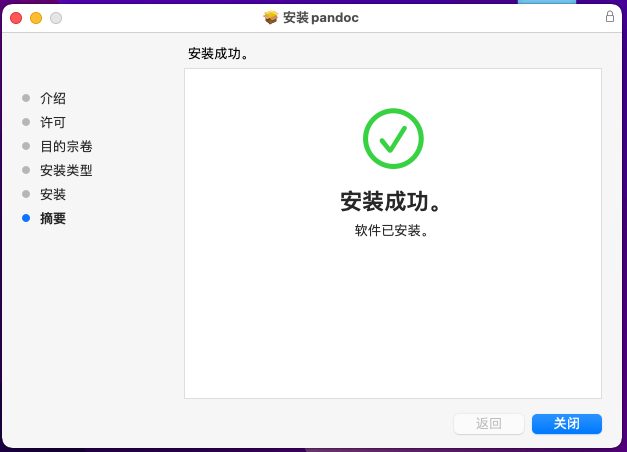
1. 点击同意



1. 点击安装



1. 关闭



# 运行

## Windows上运行

直接运行

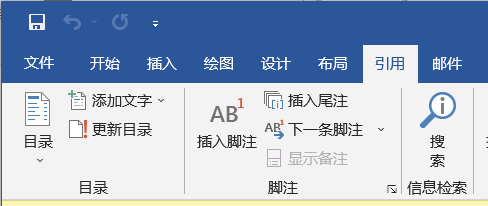
## Mac上运行

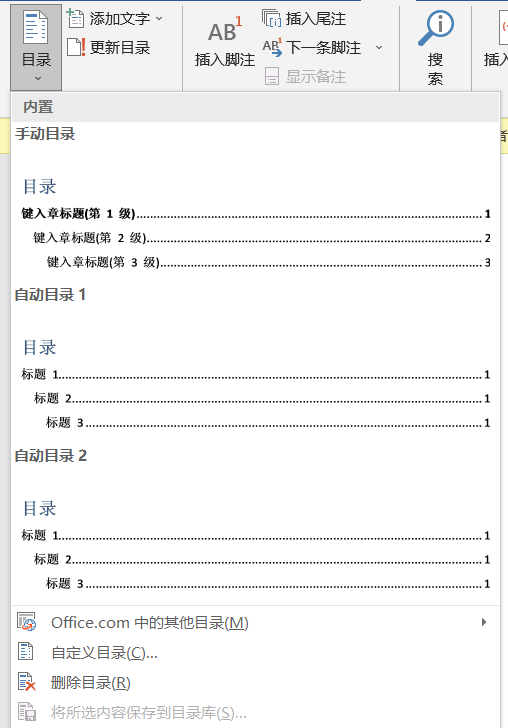
# 文档优化

## 目录的优化：

在开发过程中，我们发现docx文档中原有目录的样式会影响到转换后的匹配度，为了能够令转换的准确度变高，建议在文档编写完成后，采用自动生成目录得到文档目录。

### 采用自动生成目录





按照自动目录的形式生成目录，会得到以下形式的目录



在转化为Markdown文件后，会有如下的形式：

目录标签：



这种方式形式的目录是工具能够处理的。

## 表格的优化

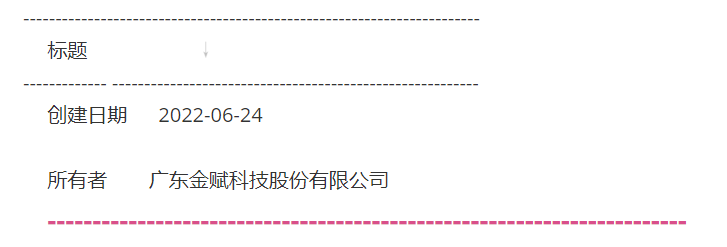
在开发过程中，我们发现docx文档中表格转换成markdown文件后，会出现两种形式。

### 转换表格类型一

文档中表格：



转换后形式：

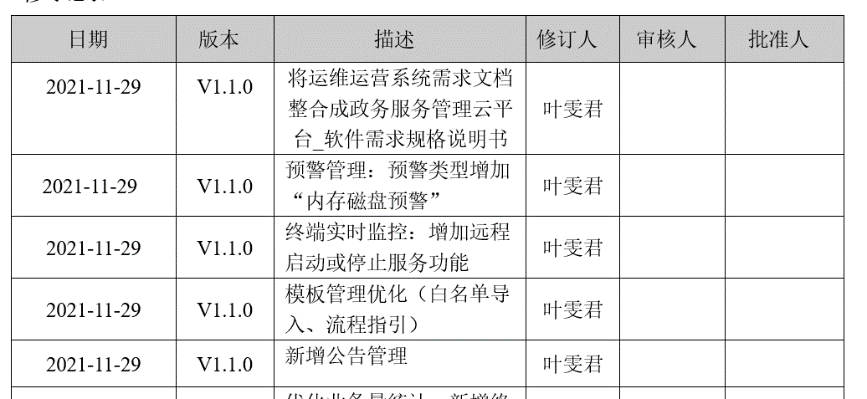


出现的条件：当表格中各单元格文本能在格内一行写完，未在格内出现换行情况，则会出现这种转换情况，由于这类型表格十分常见，在这里工具也实现了对这种表格类型的部分处理。

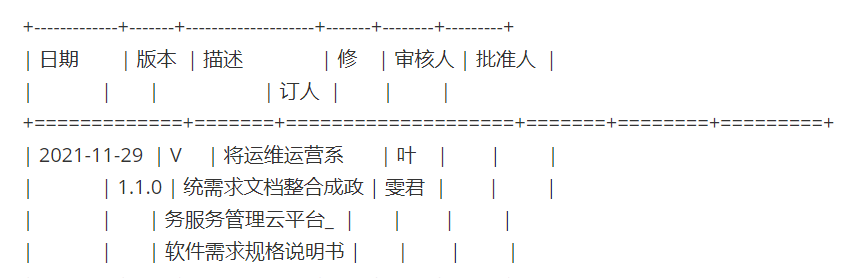
满足处理的条件是：单元格内文本未出现空格，除末列单元格外，其余单元格不能有空。

### 转换表格类型二

文档中表格;



转换后形式：



出现的条件：当表格中部分单元格文本出现格内换行情况，则会出现这种转换情况，这种形式的转换能够十分清晰的知道各单元格内文本内容。

满足处理的条件是：单元格内文本未出现空格，部分单元格格内文本出现换行。

### 处理复杂表格

当出现这种表个情况时，应该如何处理？

转换前：

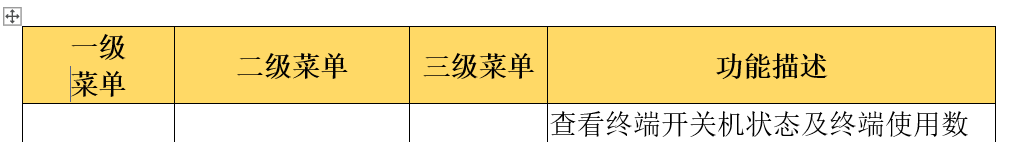


转换后：

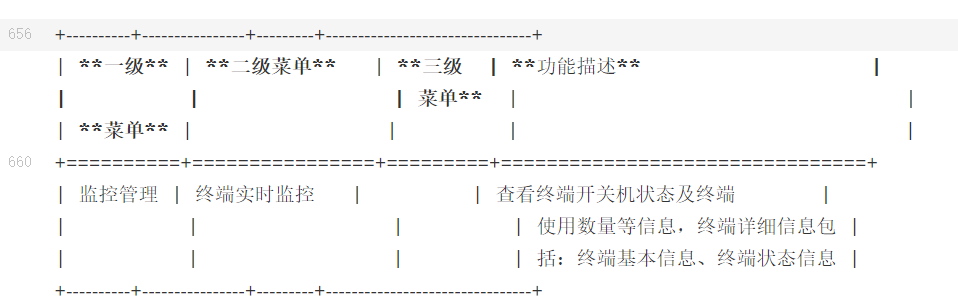


情况分析：由于markdown文件中不存在单元格合并这种表格格式，因此会出现这种十分混乱的转换，并且这种表格中有空白单元格转换工具无法正确还原出原有表格。

解决方法：此时可以在表格第一行中的单元格内人工进行换行操作。



转换后：



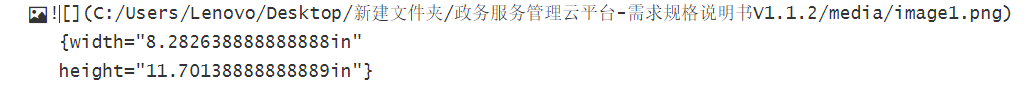


当然，先前改动的第一行的单元格内内容也需要人工更正，但这比起整个复杂表格而言，是微不足道的。

## 图片的优化

图片格式一般为常见的.png,.jpeg,.jpg;但开发中也会发现会有.emf格式的存在，这两大类图片格式在转换后的区别在与宽高参数的有无。

.png,.jpeg,.jpg



.emf



因此，为保证图片能准确转换成可上传到开发者社区的形式，需要保证当前要转换文档所在的路径上没有空格、括号、以及特殊字符。

## 其他优化

在撰写文档时，要习惯回车后写新行的内容，不要指望于wps中的自动换行，因为视觉上已经完成了换行，但本质上还处在同一行中，这样会给转换带来一些未知的因素。

# Q&A