推荐一款非常不错的从非结构化文档提取数据到excel的软件，打开office-tool.com网站 ,点“Extract Data From Unstructured Documents To Excel”选项卡

。它可以一键批量将所有非结构化文档，包括word，PDF，txt，PPT，按照关键字提取为结构化文档，也就是excel文档。

这个软件可以逐个读取所有的文档并按照配置的关键字截取每个文档中对应的内容保存在一个excel中，点击下载我们得到一个压缩包，解压后无需安装，

，直接运行Extract Data From Unstructured Documents To Excel.exe 程序即可。为了方便演示，我准备了一个文件夹，这个文件夹里面有一大堆的非结构化文档。

我打开看一看，是一些合同文档，并且还有子文件夹，这里面有很多PPT、PDF、word文档，

然后我们想要从这些合同文档提取出公司名称、中标人、评委、监察、法律顾 提取出来保存在excel中。

好的，第一步选择这些文档所在的文件夹

。第二步这里保持默认就可以，这是选择要提取那些类型的文档，你可以勾选你想提取的文档类型，

第三步 配置要生成的excel的列名称和对应的用于截取文本的前后标志，这里我们想要提取公司名称，我可以在第一列输列名称，第二列输入公司名称前的关键字：Company，第三列输入公司名称之后的关键字：一个句号，这样程序在读取每个非机构化文档时就会读取从Company开到一个句号结尾之间的内容当做公司名称填入到excel 的公司列中

同样的，我想要提取中标人，前关键字为：中标人，后关键字为：回车换行，可以用\r代表回车换行。

我们也可以空着后关键字而填写截取长度，这样程序将从前关键字开始截取指定长度的内容。

下面我们依次评委、监察、法律顾，

好，第四步选择将生成的excel保存到哪里，默认是生成到桌面，然后我们直接点开始，

此时软件会遍历此文件夹下所有的非结构化文档，然后按照这里配置的这些列去逐个的解析这些文档并提取出对应的数据到excel文件，

好，提取完毕了，桌面上已经生成了一个文excel文件，我打开看一下。

可以看到软件已经批量把每个文档读取出了公司名称、中标人、评委、监察、法律顾 并填入到此excel中了，一个文件在excel中就是一行。

好，这个软件非常方便，大家可以到office-tool.com这个网站上下载。

Hello I’ll show you an excellent software that can extracts data from unstructured documents into Excel. open website office-tool.com and click "Extract Data From Unstructured Documents To Excel" tab. This software can batch convert all unstructured documents, including Word, PDF, txt, and PowerPint, into structured document: Excel file.

This software can read every unstructured document, and extract the data according to the configured keywords, and saves that data in an Excel file.

Click the download button, you'll get a compressed package. Extract it, Then run the "Extract Data From Unstructured Documents To Excel.exe" program without installation.

I've prepared a folder that contains a large number of unstructured documents, it is some contract files, In the subfolders there are some PPT, PDF, and Word documents.

For example I want to extract company names, bid winners, judges, supervisors, and legal consultants from these contract documents and save them in an Excel file.

OK

1. select the folder where these documents are located.
2. Leave the second step as default. This allows you to choose which document types you want to extract data from. You can check the types you want to extract,
3. Configure the column names in the generated Excel and the corresponding before-and-after keywords for text extraction.

I want to extract company names, enter the column name in the first column, " Seller: " as the before-keyword in the second column, and “a company”as the after-keyword in the third column. The software will read the content between "Company" and the period as the company name and put it in the Excel's company column.  
Similarly, for extracting the bid winner, you can use " Delivery Address: " as the before-keyword and "\r" (back slash line break) as the after-keyword. Alternatively, you can leave the after-keyword blank and specify a length for extraction.

In the same way ,we can configure judges, supervisors, and legal consultants column.

1. Choose the location to save the generated Excel file. By default, it will be saved on your desktop.
2. OK that ‘s all , click "Start."

At this point, the software will traverse all the documents in the folder, analyze them based on the configured columns, and extract the corresponding data into an Excel file.

OK finished,we'll find an Excel file on the desktop. open it, you'll see that the software has read out the company name, winning bidder, judges, supervisors, and legal advisors for each document in bulk and filled them into this Excel file. Each file is one line here..

This software is incredibly convenient. You can download it from office-tool.com.