**流水/账单文件的分类需求**

**需求一：文件分类处理&去除水印**

**1、**输入的文件格式：excel、文本PDF、图像PDF

**2、功能说明：**

1）文件分类：区分文件类型；

2）去除电子印章：有印章图层的文本，去除印章图层；

3）读取数字证书信息（如果能做就做，做不了就放着）

**3、输出参数说明：**

|  |  |  |
| --- | --- | --- |
| **字段** | **说明** | **输出内容** |
| success | 执行是否成功 | true：执行成功  false：执行失败 |
| code | 报错码 | 402：无效文件 |
| data.agent\_type | 发布机构类型 | bank：银行  wechat：微信  alipay：支付宝 |
| data.file\_format | 文件格式 | excel：excel格式  TextPDF：文本PDF文件  ImagePDF：图像PDF文件 |
| data.stamp\_layer | 印章图层是否已处理 | y：有印章图层  n：没有印章图层 |
| data.digital\_certificate | 是否有数字证书 | y：包含数字证书  n：不含数字证书 |
| data.imformation | 数字证书信息 | 持有人，颁发者、是否篡改； |
| data.image | 印章图片 | base64形式 |
| flie | 去除电子印章后的文件，保持原来的文件格式； |  |

**4、交付说明：**

1）源码交付，代码里面包含dockfile

2）部署文档；

3）接口文档；