

绘到的文本数据 data

1) res1 输出。表外数据。

res1 数据是指 (表外数

(doc_type 值为 image 的是包含坐标信息的 excel 和

0 值是指位置在表格之前的表外数据, 1 值是指位置在表格之后的表外数据。

3) res2 输出 表格内数据。

包含页码和每个单元格的

) page_sum 输出, 代表百面数.

显示一共识别处理的页数.

doc_type: 代表数据来源

doc_type 为文件类型, 分为



三. 需求说明

需求 1: 从 res1 数据中，提取以下字段信息，输出用 **oures1** 字段表示；

字段	说明	输出内容
agent_type	3 种机构类型：bank（银行），alipay（支付宝），wechat（微信支付）；	bank
bank_name	如果机构类型是 alipay\wechat 的，该字段为空；银行机构名称参考“金融机构信息对照表”，输出内容包含编号和名称；	GY0001 交通银行
account_num	银行账号号码、支付宝账号、微信账号	6228480329572520271
account_name	账号户名、支付宝账号姓名、微信账号姓名	李数
idcard_num	用户身份证号码，机构类型是 alipay\wechat 的，有这个字段，bank 的没有；	330724197805240011
date_start	起始日期	2022/05/22
date_end	结束日期	2022/11/22
account_type	账户类型	对私

需求 2: 分析 res2 数据，输出**每行数据**的**行标签（row_label）**和**会计分录标签（account_label）**；输出以 **oures2** 字段表示；

1) 行标签 row_label 包含的参数：

序号	字段名	说明	输出内容
1	header_row	表头行标签，“true”表示该行是表头行，“false”表示该行不是表头行；	true
2	trader_nature	交易对象的行业分类	8850 博物馆
3	in_spend_type	收支类型，针对当前发生账户的资金方向，分为收入和支出；如果流水当中的收支类型为“其他”，则需要根据当前行的数据，判断是收入还是支出；	收入

2) 会计分录 account_label 的详细参数

序号	字段名	说明	内容
1	N-accounting_item	用来表示当前发生账户，主科目标签；参考《会计分录&科目标签-会计科目》；输出内容为编号+名称的形式；	Z0701 账户余额
2	N-accounting_entry	主科目的分录方向，分为减少和增加两	减少

		种；	
3	accounting_item	副科目标签，用来表示资金用途和交易对象；	Z0201 理财产品
4	accounting_entry	副科目的分录方向，分为减少和增加两种；	增加

需求 3：整理原数据，以规范标准格式输出原数据，输出内容以 outres3 字段表示；

机构类型 (agent_type)	字段名	格式	处理说明
bank	交易日期	2022/03/15 (yyyy/mm/dd)	如果有多个，只选择一个；
	收支类型	收入 (文本)	如果有两列表示的，合并成一列；
	交易金额	3571.42 (数值)	如果有带正负号的，则拆分成收支类型和交易金额两列；
	余额	3571.42 (数值)	
	备注 (摘要)	跨行转账 (文本)	备注如果有多个，就输出多个，以备注 1、备注 2、备注 3 形式；
	对方账号	622841.....8978 (文本)	
	对方户名	张** (文本)	
wechat	交易日期	2022/03/15 (yyyy/mm/dd)	
	交易类型	商户消费 (文本)	
	收/支/其他	收入 (文本)	
	交易方式	中国银行 (6633) (文本)	
	交易金额	3571.42 (数值)	
	交易对方	王** (文本)	
alipay	收/支	其他 (文本)	(alipay 的格式也不一定是不变的,根据原数据显示列,针对日期和金额格式做规范化处理)
	交易对方	天弘基金管理有限公司 (文本)	明细证明中“资金渠道”对应“交易对方”
	商品说明	转账收款到余额宝 (文本)	
	收/付款方式	余额宝 (文本)	明细证明中, 对应字段直接写“支付宝余额”
	金额	1.77 (数值)	明细证明中, 需要对“收入”“支出”合并;
	交易日期	2022/03/16 (yyyy/mm/dd)	

四、输出格式

参数名	说明	输出内容
success	是否正常执行成功; true: 代表执行成功 false: 代表执行失败	true
error_code	返回错误码 401: 文件格式不符合要求 402: 超过文件大小限制 -1: 未知错误	401

	工程师添加其他错误码	
message	报错信息，与 error-code 对应	文件格式不符合要求；
data	识别成功之后的数据，包含 page_sum，doc_type，outres1，outres2，outres3 数据；	

```
{
  "success": true,
  "error_code": "",
  "message": "",
  "data": {
    "page_sum": 83,
    "doc_type": "epdf",
    "outres1": {
      "agent_type": "alipay",
      "account_num": "2504767094@qq.com",
      "account_name": "王克",
      "idcard_num": "370213199907186016",
      "account_type": "对私",
      "begin_date": "2022/05/22",
      "end_date": "2022/11/22"
    },
    "outres2": [
      {
        "row_order": "001_1",
        "1.1": "收/支",
        "1.2": "交易对方",
        "1.3": "商品说明",
        "1.4": "收/付款方式",
        "1.5": "金额",
        "1.6": "交易订单号",
        "1.7": "商家订单号",
        "1.8": "交易时间",
        "row_label": {
          "header_row": true,
          "trader_nature": "",
          "in_spend_type": ""
        }
      },
      {
        "row_order": "001_2",
        "2.1": "其他",
        "2.2": "浙江网商银行股份有限公司",
        "2.3": "账户结息- 2022.11.22",
        "2.4": "",
        "2.5": "0.03",
        "2.6": "20221122102553010048250 070316309",

```

```

"2.7": "20221122101207120700000 00000670001002013INTEREST",
"2.8": "2022-11-22 15:30:43",
"row_label": {
  "header_row": false,
  "trader_nature": "商业银行服务",
  "in_spend_type": "收入",
  "account_label": {
    "N-accounting_item": "Z0701 账户余额",
    "N-accounting_item": "增加",
    "accounting_item": "S0201 营业外收入",
    "accounting_entry": "增加"
  }
},
],
"outres3": [
{
  "row_order": "001_1",
  "header_row": true,
  "1.1": "收/支",
  "1.2": "交易对方",
  "1.3": "商品说明",
  "1.4": "收/付款方式",
  "1.5": "金额",
  "1.6": "交易日期"
},
{
  "row_order": "001_2",
  "header_row": false,
  "2.1": "其他",
  "2.2": "浙江网商银行股份有限公司",
  "2.3": "账户结息-2022.11.22",
  "2.4": "",
  "2.5": 0.03,
  "2.6": "2022/11/22"
}
]
}

```

五、交付内容

- 1、源代码（包含 dockfile）
- 2、镜像文件
- 3、部署文档（包含分别采用 GPU 或 CPU 启动方式的说明）
- 4、接口文档