



<搜象 OCR 增值税电子发票结构化识别> 接口开发参考

搜象（厦门）信息技术有限公司

2022 年 02 月 21 日

目 录

1	第一章. 引擎接口开发	3
1.1	接口概述	3
1.2	服务请求报文	3
1.2.1	请求报文示例.....	3
1.2.2	请求数据项说明.....	3
1.3	服务返回报文	3
1.3.1	返回报文示例.....	3
1.3.2	返回数据项说明.....	8
1.4	失败返回码定义	8
1.4.1	失败返回示例.....	8
1.4.2	失败返回码定义.....	8
1.5	附-识别字段.....	9
1.5.1	增值税电子发票.....	9

1 第一章. 引擎接口开发

1.1 接口概述

请求地址：请参考商品页 API 接口调用地址

请求方式：POST

请求格式：Content-Type: application/json

返回类型：JSON

1.2 服务请求报文

1.2.1 请求报文示例

```
{
  "fileType": 1,
  "filePath": "http://192.168.0.2:8080/test.jpg"
}
```

1.2.2 请求数据项说明

数据项	数据类型	描述
fileType	int	识别文件类型，1-图片，2-PDF，不传则自动判断
filePath	string	识别文件路径（filePath 和 fileBase64 二者择一传参即可）
fileBase64	string	识别文件 Base64 编码（需带 Base64 头，如：图片【data:image/jpeg;base64,】, PDF【data:application/pdf;base64,】）
pageIndex	int	识别 PDF 第几页，不传则默认首页，首页从 0 开始，该参数对 PDF 有效

1.3 服务返回报文

1.3.1 返回报文示例

```
{
  "code": 0,
```

```

"data": {
  "angle": 0.12138918,
  "anglePic": 0,
  "columnResultList": [
    {
      "columnBox": [736,70,913,70,913,111,736,111],
      "columnKey": "invoice_type",
      "columnName": "发票类型",
      "columnValue": "普通发票"
    },
    {
      "columnBox": [1054,66,1177,66,1177,87,1054,87],
      "columnKey": "invoice_code",
      "columnName": "发票代码",
      "columnValue": "033002100411"
    },
    {
      "columnBox": [1054,98,1141,98,1141,120,1055,120],
      "columnKey": "invoice_number",
      "columnName": "发票号码",
      "columnValue": "17928286"
    },
    {
      "columnBox": [1057,135,1202,135,1202,153,1057,153],
      "columnKey": "invoice_date",
      "columnName": "开票日期",
      "columnValue": "2021-08-06"
    },
    {
      "columnBox": [1057,166,1267,166,1267,187,1057,188],
      "columnKey": "check_code",
      "columnName": "校验码",
      "columnValue": "44544284273153244679"
    },
    {
      "columnBox": [126,169,252,169,252,193,126,193],
      "columnKey": "machine_code",
      "columnName": "机器编号",
      "columnValue": "661836820867"
    },
    {
      "columnBox": [221,205,519,204,519,235,221,236],
      "columnKey": "buyer_name",

```

```

        "columnName": "购买方名称",
        "columnValue": "搜象（厦门）信息技术有限公司"
    },
    {
        "columnBox": [221,239,524,238,524,265,221,266],
        "columnKey": "buyer_Identifier",
        "columnName": "购买方纳税人识别号",
        "columnValue": "91350211MA33G2XB85"
    },
    {
        "columnBox": [844,203,1290,202,1291,328,844,328],
        "columnKey": "password_area",
        "columnName": "密码区",
        "columnValue":
"+72<70139<58<109774*/5+3<0\n/+ -9<- +<9*55591</<75+5+8-*8\n3////327>5867155>90<
<65*/82\n-228-899+2*/-*9*581*3+>336-"
    },
    {
        "columnBox": [36,376,261,376,261,407,36,407],
        "columnKey": "project_col",
        "columnName": "货物或服务名称列",
        "columnValue": "*计算机外部设备*打印机"
    },
    {
        "columnBox": [562,376,589,376,589,405,562,405],
        "columnKey": "unit_col",
        "columnName": "单位列",
        "columnValue": "台"
    },
    {
        "columnBox": [759,378,875,378,875,403,759,403],
        "columnKey": "unit_price_col",
        "columnName": "单价列",
        "columnValue": "276.99115044"
    },
    {
        "columnBox": [1010,379,1068,379,1068,402,1010,402],
        "columnKey": "amount_col",
        "columnName": "金额列",
        "columnValue": "276.99"
    },
    {
        "columnBox": [1094,378,1136,378,1136,402,1095,402],

```

```
        "columnKey": "tax_rate_col",
        "columnName": "税率列",
        "columnValue": "13%"
    },
    {
        "columnBox": [1286,377,1338,377,1338,405,1286,405],
        "columnKey": "tax_amount_col",
        "columnName": "税额列",
        "columnValue": "36.01"
    },
    {
        "columnBox": [959,600,1069,600,1069,624,960,624],
        "columnKey": "total_amount",
        "columnName": "合计金额(元)",
        "columnValue": "276.99"
    },
    {
        "columnBox": [1246,598,1336,598,1336,627,1246,627],
        "columnKey": "total_tax_amount",
        "columnName": "合计税额(元)",
        "columnValue": "36.01"
    },
    {
        "columnBox": [404,640,551,640,551,668,404,668],
        "columnKey": "total_amount_tax_cap",
        "columnName": "价税合计(大写)",
        "columnValue": "叁佰壹拾叁圆整"
    },
    {
        "columnBox": [1098,645,1212,645,1212,666,1098,666],
        "columnKey": "total_amount_tax_low",
        "columnName": "价税合计(小写/元)",
        "columnValue": "313.00"
    },
    {
        "columnBox": [221,683,457,684,457,712,221,711],
        "columnKey": "seller_name",
        "columnName": "销售方名称",
        "columnValue": "杭州天赐成电脑有限公司"
    },
    {
        "columnBox": [221,712,526,713,526,742,221,741],
        "columnKey": "seller_Identifier",
```

```
        "columnName": "销售方纳税人识别号",
        "columnValue": "91330106749454776Y"
    },
    {
        "columnBox": [221,745,646,746,646,775,221,774],
        "columnKey": "seller_addr_tell",
        "columnName": "销售方地址电话",
        "columnValue": "西湖区文三路 251 号欣荣大厦 B 幢 604 室 56776828"
    },
    {
        "columnBox": [221,780,550,781,550,805,221,804],
        "columnKey": "seller_bank_account",
        "columnName": "销售方开户行及账号",
        "columnValue": "建行高新支行 33001616735050002080"
    },
    {
        "columnBox": [160,817,224,817,224,842,160,842],
        "columnKey": "cashier",
        "columnName": "收款人",
        "columnValue": "陈娇丽"
    },
    {
        "columnBox": [471,816,547,816,547,844,471,844],
        "columnKey": "reviewer",
        "columnName": "复核",
        "columnValue": "汪志芬"
    },
    {
        "columnBox": [784,816,853,816,853,841,784,841],
        "columnKey": "drawer",
        "columnName": "开票人",
        "columnValue": "王保平"
    }
],
"templateCatelId": "825454941819633664",
"templateCateName": "增值税发票",
"templateId": "911194901230649344",
"templateName": "增值税电子发票"
},
"message": "success"
}
```

1.3.2 返回数据项说明

数据项	数据类型	描述
code	int	返回编码，0-代表成功，其余失败
data	object	返回具体数据项
+angle	float	文字顺时针偏转方向
+anglePic	int	图片四分类方向，0-正向, 90-右向, 180-反向, 270-左向
+templateCateId	string	模板分类 ID
+templateCateName	string	模板分类名称
+templateId	string	模板 ID
+templateName	string	模板名称
+columnResultList	objectArray	文字识别数据项列表
++columnName	string	字段中文名称
++columnKey	string	字段英文名称
++columnValue	string	字段值（\n 字符表示字段值中的换行或较大间隔）
++columnBox	intArray	字段坐标->[x1,y1,x2,y2,x3,y3,x4,y4]（非固定项，坐标方式提取的字段有返回，规则方式提取的字段无返回）
message	string	返回信息，包括成功失败信息

1.4 失败返回码定义

1.4.1 失败返回示例

```
{
  "code": 1011,
  "data": {},
  "message": "无识别文件"
}
```

1.4.2 失败返回码定义

返回编码	返回信息	描述
1011	无识别文件	filePath 和 fileBase64 同时为空
1012	不支持的文件类型	支持图片和 PDF，图片包括但不限于 jpg、jpeg、png、

		bmp、gif、tif、tiff、jfif、heic、jpe、dib 等格式
1014	文件 URL 错误	文件 URL 有误，读取不到资源
1015	文件 BASE64 串错误	文件 BASE64 串有误，解析文件异常
1099	服务器端异常	返回具体异常信息

1.5 附-识别字段

1.5.1 增值税电子发票

字段中文名称	字段英文名称	字段中文名称	字段英文名称
发票类型	invoice_type	单价列	unit_price_col
发票代码	invoice_code	金额列	amount_col
发票号码	invoice_number	税率列	tax_rate_col
开票日期	invoice_date	税额列	tax_amount_col
校验码	check_code	合计金额(元)	total_amount
机器编号	machine_code	合计税额(元)	total_tax_amount
购买方名称	buyer_name	价税合计(大写)	total_amount_tax_cap
购买方纳税人识别号	buyer_identifier	价税合计(小写/元)	total_amount_tax_low
购买方地址电话	buyer_addr_tell	销售方名称	seller_name
购买方开户行及账号	buyer_bank_account	销售方纳税人识别号	seller_identifier
密码区	password_area	销售方地址电话	seller_addr_tell
货物或服务名称列	project_col	销售方开户行及账号	seller_bank_account
规格型号列	model_col	收款人	cashier
单位列	unit_col	复核	reviewer
数量列	number_col	开票人	drawer