

# hjxh\_express\_match 皇家小虎快递匹配系统

---

## 1. Diary

1. [2021-12-30](#)
2. [2021-12-29](#)
3. [2021-12-28](#)
4. [2021-12-26](#)
5. [2021-12-24](#)

## 2. Todo

1.  优化文件读取过程中的前端界面展示
2.  测试ERP的数据读取与存储
3.  实现TRD的数据读取与存储
4.  实现TRD与ERP的对比
5.  TS2339: Property 'erp' does not exist on type 'PrismaClient' .
6.  solve named capture problem of webstorm
7.  研究.d.ts

## 3. Finished/Bugfix

1.  解析数据有误问题
2.  sqlite3、typeorm native dependency
3.  学习prisma的连接与插入业务流设计范式。
4.  csv只读取第一行
5.  寻求fast-csv skip error的方案。
6.  支持上传的文件的 Sample 备份预览（包含head与tail五行）。
7.  完成数据库、前端、后端的基本设计。2021年12月22日
8.  no-headers pass but headers not
9.  progress-stream cause bug
10.  try...catch...finally problem
11.  ipcRenderer duplicate response
12.  how to asynchronously and partially read csv
13.  which to choose: node-csv or fast-csv
14.  axios form-data parse bug

## 4. 业务问题

1.  ERP表中，计算价格错误
2.  第三方表匹配不上ERP表
3.  第三方报表格式
4.  第三方表字段含义

## 5. 设计

1. 工作流设计
2. 数据库的选择
3. 数据库的表设计
4. [DEPRECIATED] Module Design

## 6. 心得

1. Interface helps do right!
2. Interface 和 Object 之间的关系
3. Error类的继承设计

## 7. [ARCHIVE] deploy script

1. deploy backend (python, fastapi)
2. deploy frontend (node, react)

## 8. [ARCHIVE] 表约定【重要】

1. 术语定义
2. 通用表约定
3. erp表约定
4. trd表约定
5. 列字段约束

# Diary

2021-12-30

- ★ 升级了程序的整体结构，奠定了前后端通信范式，整个项目开始规范化、模块化
- 终于找到了csv读取quote异常导致中断的解决方案
- 实现了对csv读取的再一次抽象封装

2021-12-29

完成了数据库ORM的选型工作，确定使用Prisma

2021-12-28

前端UI基本成型

2021-12-26

深入理解并实现了markdown的自定义解析

2021-12-24

完成了系统前端与业务流的重构

# Todo

- ➊ 优化文件读取过程中的前端界面展示
- ➋ 测试ERP的数据读取与存储
- ➌ 实现TRD的数据读取与存储
- ➍ 实现TRD与ERP的对比
- ➎ TS2339: Property 'erp' does not exist on type 'PrismaClient' !

## 解决 named capture 问题的 WebStorm

文件路径: [[erb/src/main/erp.ts]]

```
export const parseDate = (dateInput: string): string => {
  let g: RegExpMatchArray | null = dateInput.match(
    matcher: /(?<year>20\d\d)[\s](?<month>\d\d)[\s](?<day>\d\d)/
  );
  if (!g) {
    g = dateInput.match(matcher: /(?<month>\d\d)\|(?<day>\d\d)\|(?<year>20\d\d)/);
    // prettier-ignore
    if (!g) throw new MyError({ errorType: ValidateError.Invalid, msg: `should date ${dateInput} satisfy 'YYYY-MM-DD' format` });
  }
}
```

## 研究.d.ts

### 完成/Bugfix

#### 解析数据有误问题

```
reading file, name: /Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/data/2021-11/erp/erp_20211214.销售出库单_李佳欣_1.csv
{ encoding: 'utf-8' }
Error: Parse Error: expected: ',' OR new line got: '九'. at '九江市武宁县江西省九
at QuotedColumnParser.checkForMalformedColumn (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/column/QuotedColumnParser.ts:100:23)
at QuotedColumnParser.parse (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/column/QuotedColumnParser.ts:42:16) ...
at ColumnParser.parse (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/parser/column/ColumnParser.ts:24:44)
at RowParser.parse (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/parser/RowParser.ts:45:48)
at Parser.parseRow (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/parser/Parser.ts:76:36)
at Parser.parseWithoutComments (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/parser/Parser.ts:46:35)
at Parser.parse (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/parser/Parser.ts:39:21)
at CsvParserStream.parse (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/CsvParserStream.ts:109:44)
at CsvParserStream._transform (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/CsvParserStream.ts:83:31)
at CsvParserStream.Transform._write (/Users/mark/Documents/mark_projects/皇家老虎/HJXH/hjxh_express_match/erb/node_modules/@fast-csv/parse/src/CsvParserStream.ts:384:23)
```

#### WebStorm CSV

CSV file content:

	王云飞,1.0000,1,9bhv\$^~杨*,普通仓库,盛世星城*栋*单元*室,15*****
1, vTouEZkZ\$^~张*士,普通仓库,平川大道银盛**,13*****51,,,网店销	
1, 0/H3YKvk\$^~罗*芳,普通仓库,"九江市武宁县江西省九江市武宁县*****	
1, eBfawwN\$^~吴*花,普通仓库,燕山台板*,15*****55,,,网店销售,0.	
1, pVbPwZFP\$^~王*汉,普通仓库,焦祈*,17*****85,,,网店销售,0.00,0	
1, NwPPYoMS\$^~黄*平,普通仓库,工农路*号大拇指幼儿园,18*****59,,	
1, aF0E\$^~陈*,普通仓库,中建国际花园一期G*号楼(槐荫公园北),13*****	

Terminal output:

```
'拣货员',          '打包员',          '验货员',          '检视员',
'发货员',          '打印波次',          '物流单打印状态',  '发货单打印状态',
'分拣单打印状态', '支付时间',          '分拣单编号',        '发货倒计时',
'外部单号',          '签出人',          '包装',          '处理天数',
'审核创建时间',      '拦截原因',          '打印备注',        '体积',
'分销商',          ''
```

Error: Parse Error: expected: ',' OR new line got: '九'. at '九江市武宁县江西省九

## wps csv

The screenshot shows a detailed data entry interface. The main table has columns labeled E through Z. The first few rows contain header-like information such as '已出库, 物款到发货' and '18 王云飞'. The table is filled with data including names like '王云飞', item codes like '160006833', and various numbers. A 'Find' dialog box is open at the top, containing the text '九江市武宁县江西省九江市武宁县\*\*\*\*\*'. The status bar at the bottom displays the full file path.

## sublime csv

This screenshot shows a large CSV file in Sublime Text. The table has multiple columns including '184289', 'JY2021110698264', '211106-229601483681141', 'CK2021110813564', '美粮坊-湖南雅集仓, 美粮兴国县, 美粮坊-长沙-多多韵达, 2.0000,' and so on. The file appears to be a log of transactions or inventory management.

## attempt to save to wps [update: 不愧是我]

A modal dialog box is centered over the WPS interface, prompting the user to decide whether to keep incompatible features. The message says: 'The file you are trying to save may contain features that are not compatible with current file format. Do you want to keep the workbook in this format?' Below this are two buttons: 'No' and 'Yes'. The background table continues from the previous screenshots, showing more rows of data.

This solution is rather robust, and does solve my problem, which achieves a balance between server code and user experience!

What a genius I am that a flash of inspiration occurred to my mind which suggested me to give 're-saving' a shot !

```
0/2021 0:00","","","","系统","无","45","11/6/2021 13:57","","","","0","",""],"progress":0.938086.901960785}],"msg":"should _id: () not null"}'
{"line":221621,"row":["JY2021110293363","2021102900017-30","CK20211104120158","7479d27bb","济南多多","普通仓库","海尔路原天山二路100号日日顺物流即墨物流产业园","打印","10/30/2021 0:00","","","","系统","无","45","11/4/2021 19:57","","","","0","",""],"speed":10868893.538461538}],"msg":"should _id: () not null"}'
{"line":225497,"row":["JY2021103122410","202110251544-27","CK2021110382768-2","7bf8","万州美团优选","普通仓库","区明品福冷链物流园","18523204458","18523204458","0:00","","","","系统","无","48","11/6/2021 13:57","","","","0","",""],"progress":{"percentage":0.76923077}],"msg":"should _id: () not null"}'
finished
(base) ➜ erb git:(main) ✘
```

## ■ sqlite3、typeorm native dependency

2021-12-30 update: there is no need to think about `typeorm` any longer since I have decided to use `prisma` which works well.

安装`sqlite3`、`typeorm`之后没法运行`electron`了

Terminal: Local +   
(base) ➜ erb git:(main) npm i

```
> postinstall
> ts-node .erb/scripts/check-native-dep.js && electron-builder install-app-deps && cross-env NODE_ENV=development TS_JSDOC_DISABLE_INDEX=true electron-rebuild
[...]
Webpack does not work with native dependencies.
sqlite3, typeorm are native dependencies and should be installed inside of the "./release/app" folder.
First, uninstall your package from "./package.json".
npm uninstall your-package
Then, instead of installing the package to the root "./package.json":
npm install your-package
Install the package to "./release/app/package.json"
cd ./release/app && npm install your-package
Read more about native dependencies at:
https://electron-react-boilerplate.js.org/docs/adding-dependencies/#module-structure

npm ERR! code 1
npm ERR! path /Users/mark/Documents/mark_projects/皇家小虎/HJXH/hjxh_express_match/erb
npm ERR! command failed
v NOD
npm ERR! command sh -c ts-node .erb/scripts/check-native-dep.js && electron-builder install-app-deps && cross-env dev.dll.ts && opencollective-postinstall

npm ERR! A complete log of this run can be found in:
npm ERR!   /Users/mark/.npm/_logs/2021-12-29T00_58_57_231Z-debug.log
[...]
```

ref:

- What is exactly native dependency? · Issue #1042 · electron-react-boilerplate/electron-react-boilerplate

## ■ 学习prisma的连接与插入业务流设计范式。

方案: `prisma`会在第一次query时自动连接数据库, 拥有一个数据库连接池, 所以无需我自己管理。ref: [Connecting and disconnecting \(Concepts\) | Prisma Docs](#)

## ■ csv只读取第一行

方案：在`fast-csv`的第一个回调里处理就可以，这里就是第一行；如果不是行的话，可以用`fs`的`start | end`参数控制。

## 寻求`fast-csv` skip error的方案。

结果：官方回复：[Skip row at on\('error'\) event](#)

Skip row at on('error') event #179

**Closed** loretoparisi opened this issue on Mar 18, 2017 · 1 comment

loretoparisi commented on Mar 18, 2017

It would be possible to skip the current parsed row when the 'error' event is fired? That would help in case of malformed rows.

doug-martin commented on Jul 11, 2019

No, if an error is encountered it is usually because the parser is unsure how to proceed with the file, and can lead to very unpredictable results.

doug-martin closed this on Jul 11, 2019

## 支持上传的文件的 Sample 备份预览（包含head与tail五行）。

这是之前的方案，目前已经不采用，目前不存在0/1的问题，即不是按文件为基本单位传输给用户，而是在读取过程中持续地按行为基本单位（可选：筛选出有问题的部分）传输给用户，因此备份预览没有意义也不需要了。

## 完成数据库、前端、后端的基本设计。2021年12月22日

### no-headers pass but headers not

The reason is that I wrote condition of `_id === null` with always return `true` since the default 'error result' is `""`, an empty string not `null`.

So, I remedied it by changing the condition to `!_id` which can detect not only `null` but also 'empty'. And then I fixed this bug.

```
export const validateErpItemOfNoHeader = (
  data: (string | number)[],
  erpPosMap: ErpPosMap
): ErpItem => {
  let _id: string | number = data[erpPosMap._id];
  if (!_id) throw new MyError({ errorType: ValidateErrorNotNull, msg: `'_id': ${_id}` });
  _id = String({ value: _id }).replace({ searchValue: /="/, replaceValue: '' });
}
```

## progress-stream cause bug

2021年12月29日04:47:04，在引入`progress-stream`包后导致了`csv`解析的错误，后来测试发现，只要把`progress-stream`放在`fast-csv`之前就行了。想来也是，毕竟`progress-stream`是处理`stream`的，而

csv那一步已经变成row了，具体细节我也不明白也很重要，这里会用就行了。

```
/**  
 * 需要把`pipeProgressStream`放在`csv`之前  
 */  
return s  
.pipe(destination: pipeProgressStream) ProgressStream  
.pipe(destination: csv.parse(args: { headers: false })) CsvParserStream<Row<any>, Row<any>>  
  .pipe(destination: fileWriter) FileWriteStream<Row<any>>
```

- freeall/progress-stream: Read the progress of a stream
- progress-stream - npm
- node.js - streams with percentage complete - Stack Overflow

## try...catch...finally problem

2021年12月29日03:40:04，try...catch结构中，finally为什么会提前结束？事实上这个问题是我对js中的try...catch...finally理解不够深刻，还拿着python中的同步思维去理解的。js里的这套结构体远比我想象地复杂，但是呢，为了避免这种复杂（在try或者catch中各种乱返回），一种好的办法就是只在finally里返回（当然，这点我是知道的，只不过没有把它当做信仰）。

更新：我后续在那块代码中修改了写法，剔除了try...catch...，所以finally的问题也就不存在了。

启示：尽量不要在异步程序中使用t....c....f，否则可要小心了。

- Finally in Promises & Try/Catch - DEV Community
- javascript - Why does a return in finally override try? - Stack Overflow

## ipcRenderer duplicate response

2021年12月29日01:03:02，ipc通信中前端逐步累积渲染问题，猜测原因，可能是1. ipcRenderer中的端口使用on导致重复监听，并且最后的removeAllListeners方法没有生效；2.react问题。貌似重新启动一下electron就好了.....

```
selecting file                                         UploadErp.tsx:53
listing listeners of                                     VM113_preload.js:23
requestReadFile
▶ []
received files return:                                UploadErp.tsx:58
{sendTime: Wed Dec 29 2021 03:35:25 GMT+0800 (China
Standard Time), content: Array(1), level: 13}
content is over                                         UploadErp.tsx:4
removing channel: requestReadFile VM113_preload.js:1
removed channel, now listeners:  VM113_preload.js:1
▶ []
>
```

更新：实际上这个问题，是由try...catch...finally使用不当导致的，和electron、react都没关系，它们的部分都是正常的，electron的listener监听部分失效，但是重新启动一下它也正常了。关于t....c....f问题见try...catch...finally problem

## █ how to asynchronously and partially read csv

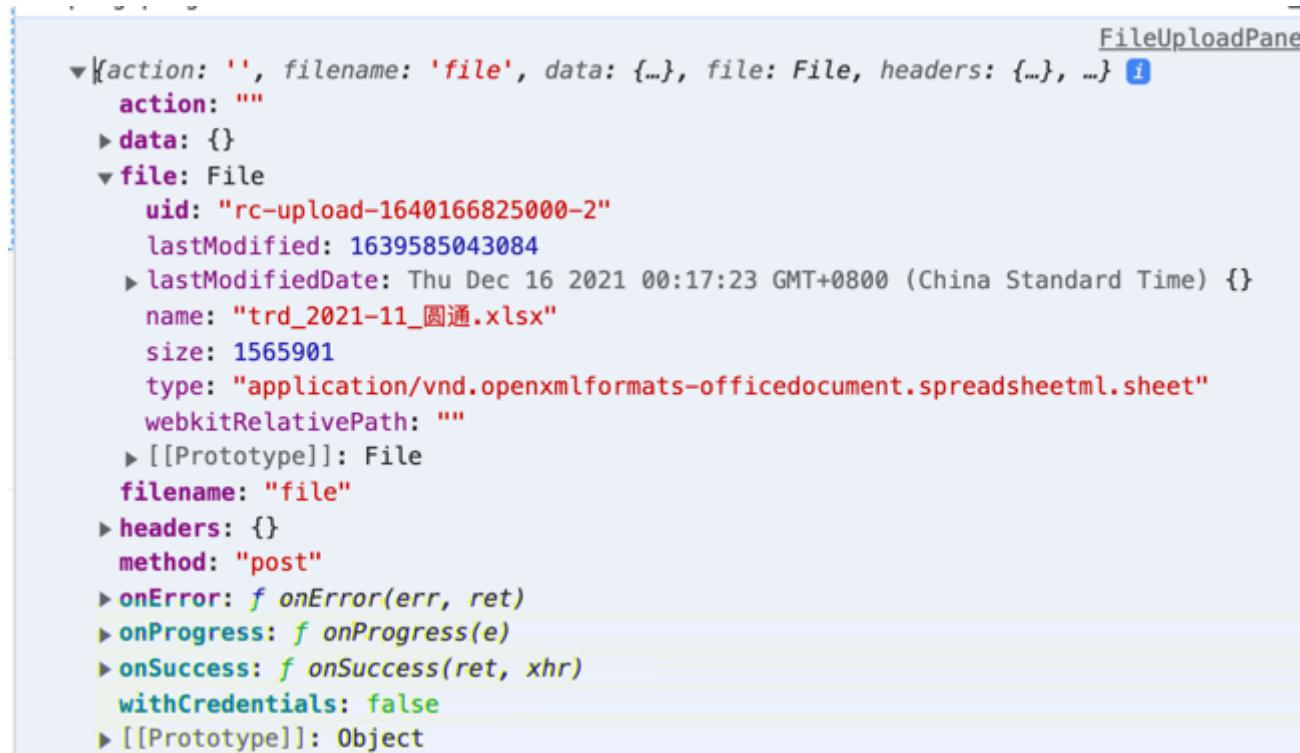
csv异步、小量快读读取.csv文件头部信息，以确定编码。已实现，基于`fast-csv`解决了中文乱码导致`node-csv`无法读取的问题，同时基于`iconv`实现了`gbk`与`utf-8`之间的无缝转换

## █ which to choose: `node-csv` or `fast-csv`

csv文件读取的选型与方法。经过鉴定，`node-csv`的接口比较低级，`fast-csv`更高些，并且更加稳健，输出方式（可以设置headers有或者无）比较友好，所以选择`fast-csv`。在读取上，有`fs.read`，`fs.readFile`，`fs.createReadStream`等几种方式，经过比较，`fs.read`接口比较低级，速度快，适合用于编码测试；等测试完后使用`fs.createReadStream`处理流数据比较好，方便与`iconv`、`fast-csv`等配合。

## █ axios `form-data` parse bug

本地前端上传文件`options`信息：



```
FileUploadPane
▼ {action: '', filename: 'file', data: {...}, file: File, headers: {...}, ...} ⓘ
  action: ""
  ▶ data: {}
  ▶ file: File
    uid: "rc-upload-1640166825000-2"
    lastModified: 1639585043084
    ▶ lastModifiedDate: Thu Dec 16 2021 00:17:23 GMT+0800 (China Standard Time) {}
    name: "trd_2021-11_圆通.xlsx"
    size: 1565901
    type: "application/vnd.openxmlformats-officedocument.spreadsheetml.sheet"
    webkitRelativePath: ""
    ▶ [[Prototype]]: File
    filename: "file"
  ▶ headers: {}
  method: "post"
  ▶ onError: f onError(err, ret)
  ▶ onProgress: f onProgress(e)
  ▶ onSuccess: f onSuccess(ret, xhr)
  withCredentials: false
  ▶ [[Prototype]]: Object
```

远程前端上传文件`options`信息：

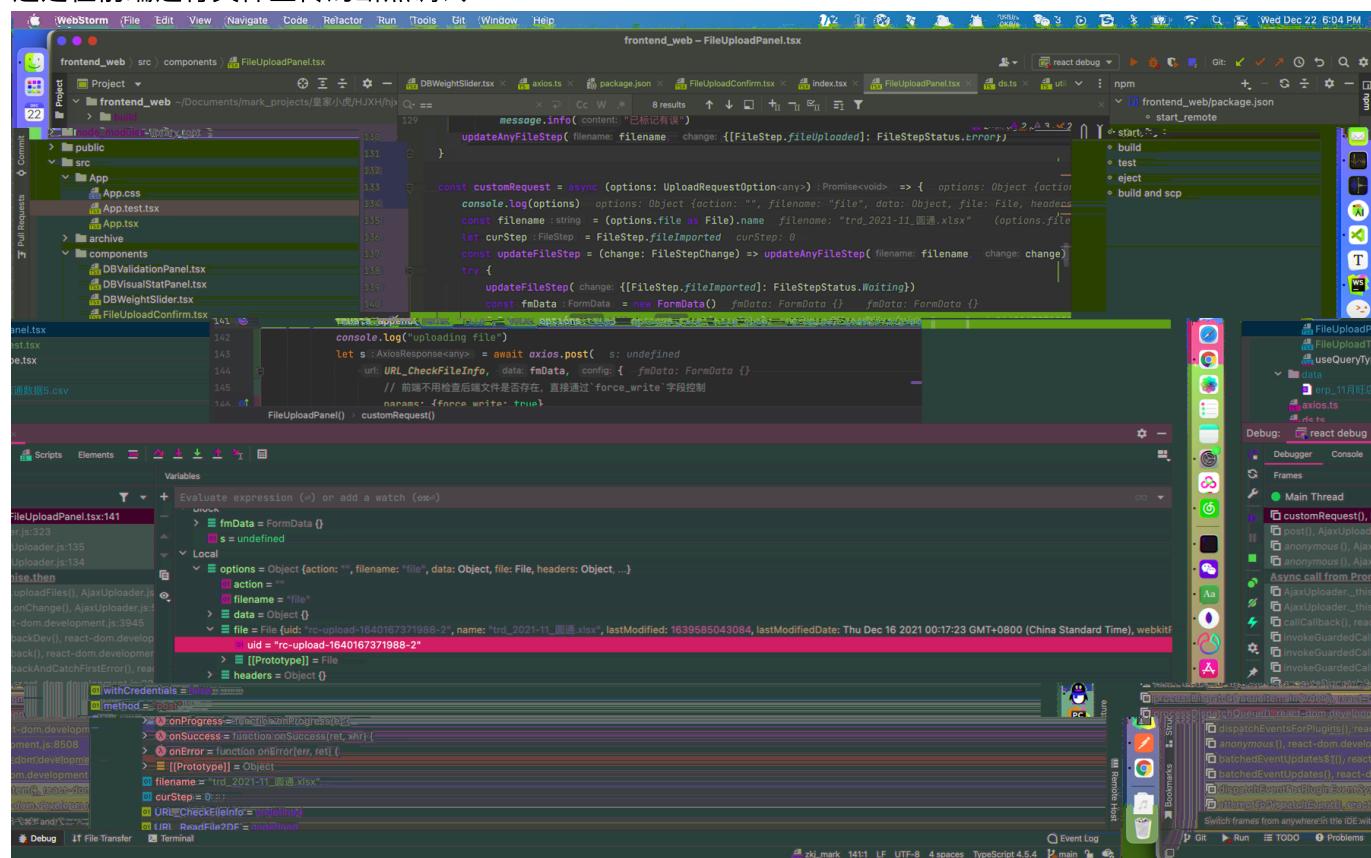
net:::ERR\_UNKNOWN\_URL\_SCHEME

FileUploadPanel.ts

```

▼ {action: '', filename: 'file', data: {...}, file: File, headers: {...}, ...} ⓘ
  action: ""
  ▶ data: {}
  ▶ file: File
    uid: "rc-upload-1640168081852-2"
    lastModified: 1639585043084
  ▶ lastModifiedDate: Thu Dec 16 2021 00:17:23 GMT+0800 (China Standard Time) {}
    name: "trd_2021-11_圆通.xlsx"
    size: 1565901
    type: "application/vnd.openxmlformats-officedocument.spreadsheetml.sheet"
    webkitRelativePath: ""
  ▶ [[Prototypell]: File
    filename: "file"
  ▶ headers: {}
    method: "post"
  ▶ onError: f onError(err, ret)
  ▶ onProgress: f onProgress(e)
  ▶ onSuccess: f onSuccess(ret, xhr)
    withCredentials: false
  ▶ [[Prototypell]: Object
  ...
  ...
  ...
  
```

这是在前端进行文件上传的断点调试



基于这个，再进行服务端文件调试，比对文件信息的不同。但是现在的问题是服务端进入不了程序逻辑，直接被fastapi拒绝了。

## 业务问题

ERP表中，计算价格错误

目前已发现的主要有两种错误：

1. 收货地区填写不规范 (6/50+w) , 导致未能正确识别省份名称, 例如:

收货地区	实际重量	物流单号	error_type
拱墅区 泰嘉园B座 407	17.68	= "432188818204088"	收货地区的省份不合法
"4301020116407007" 收货地区的省份不合法		山东省 济宁市	4.26
026028697911 收货地区的省份不合法		杭州市 杭州市 江干区	9.6933
= "432202602869888" 收货地区的省份不合法		杭州市 杭州市 江干区	6.08
54237551473" 收货地区的省份不合法		安徽省 滁州市 潘阳县	12.96
49641261123" 收货地区的省份不合法		江苏省 杭州市 江干区	17.68
(N/A)	(N/A)	(N/A)	(N/A)

更新: 对于这种问题, 直接提示报错即可。

2. 重量为0 (这个还比较多, 146/50+w) , 无法理解, 例如:

收货地区	实际重量	物流单号	error_type
辽宁省 大连市 普兰店区	0	= "432207278376727"	重量不合法
辽宁省 大连市 金州区	0	= "432207278376821"	重量不合法
辽宁省 太原市 甘井子区	0	= "432207278376915"	重量不合法
河北省 廊坊市 三河市	0	= "432207278377064"	重量不合法
四川省 成都市 成华区	0	= "432207278377159"	重量不合法
云南省 德宏傣族景颇族自治州 盈江县	0	= "432207278377254"	重量不合法
黑龙江省 哈尔滨市 平房区	0	= "432207278377349"	重量不合法
辽宁省 大连市 金州区	0	= "432207278377444"	重量不合法
黑龙江省 大庆市 让胡路区	0	= "432207278377539"	重量不合法
辽宁省 大连市 沙河口区	0	= "432207278377599"	重量不合法
辽宁省 大连市 甘井子区	0	= "432207280718895"	重量不合法

更新: 对于这个问题, 一开始是某张表给的问题, 后续小范围问题可以直接报错提示。

## ■ 第三方表匹配不上ERP表

更新: 初期是因为发现有很多表的id导出是=开头的, 后期加了稳健的检测与转换, 这个问题就基本没了。

## ■ 第三方报表格式

注意到发来的第三方对账单, 例如: "11月第三方仓韵达"是Excel格式, 且包含着"订单明细wms"表与"快递核算标准"表, 请问这个应该于"惯例"吧?

anyway, 这个倒不是啥问题哈, 个人可以接受。

更新: 这个问题其实不重要, 因为有一个专门的快递核算大表, 可以根据那个进行运费计算。

## ■ 第三方表字段含义

	发货订单号	快递公司	快递单号	所属快递	省份	订单商品总数	订单商品种类数	账单重量	发货时间	快递费	涨价金额
0	DO20211070000047874	长沙菜鸟 韵达	432169686442736	韵达快递	江西省	1	1	0.900	2021-11-07	2.6	0.3
1	DO20211070000047873	长沙菜鸟 韵达	432169686442646	韵达快递	广东省	1	1	0.900	2021-11-07	2.6	0.3
2	DO20211070000045970	长沙拼多多 韵达	462082444765915	韵达快递	广东省	2	1	1.958	2021-11-07	3.2	0.3
3	DO20211070000045968	长沙菜鸟 韵达	432169644541652	韵达快递	广东省	2	1	2.000	2021-11-07	3.2	0.3
4	DO20211070000045967	长沙抖音 韵达	3180399197731	韵达快递	江西省	2	1	2.000	2021-11-07	3.2	0.3

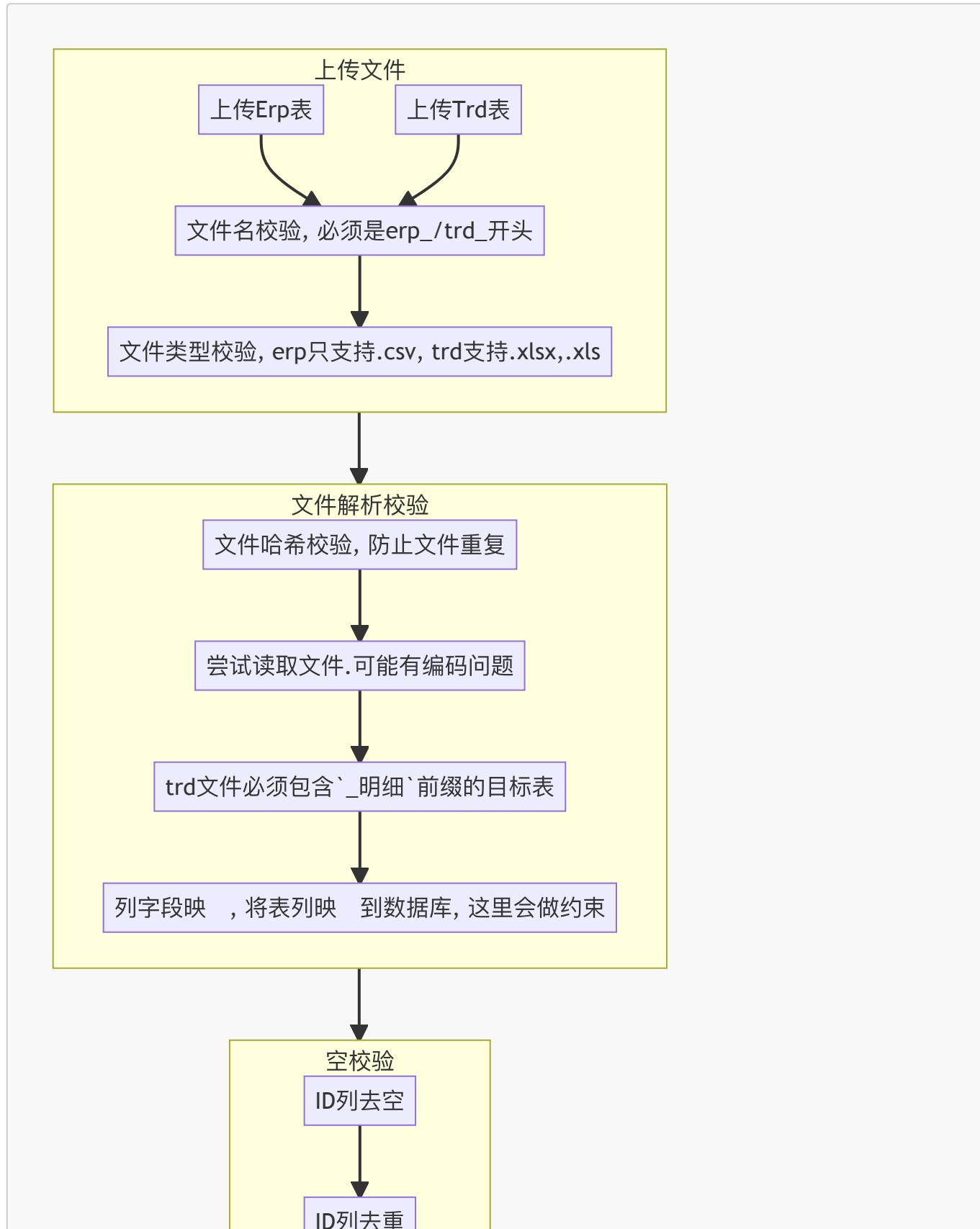
1. 发货订单号 和 快递单号 之间的区别? 该选用哪个?

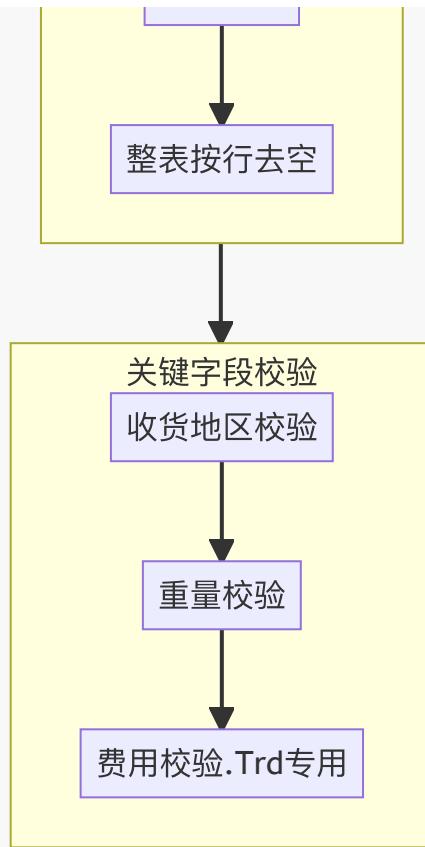
2. 涨价金额 是不是无关紧要?

更新: 目前直接规定死字段。

## 设计

### 工作流设计





## 数据库的选择

对数据库选择的倾向：mongodb --> mysql --> sqlite3

【2021年12月28日】我是觉得，应该用sqlite3，也是今天才意识到的。

## Electron app with database

Asked 3 years, 6 months ago Active 1 year, 5 months ago Viewed 27k times



I'm creating a web app for ticket reservation. The only problem is the database. I don't want to tell my client to install XAMPP or set a database, etc.

33

Is there any way to package the app with the database?

javascript sql node.js database electron

11

Share Edit Follow Flag

asked Jun 30 '18 at 23:18

gian29

610 ● 1 ● 7 ● 9

1 sqlite might be an option – David784 Jun 30 '18 at 23:21

Did you find any satisfying solution? – C.P. Oct 28 '18 at 6:52

1 check this out using sqlite with electron [github.com/sjmelia/electron-boilerplate-sqlite](https://github.com/sjmelia/electron-boilerplate-sqlite) – Nyi Nyi Ja 3 '19 at 4:40

Add a comment

Active Oldest V

## 2 Answers

30

[SQLite](#) is a good option for a local database that you can ship with your application. It won't require the user to setup or install any extra dependencies. It's certainly not as full featured as a full on server, but it should be good enough for a local desktop app.

Share Edit Follow Flag

answered Jul 1 '18 at 1:12

Jacobm001

4,180 ● 4 ● 30 ● 50

ref:

- [javascript – Electron app with database – Stack Overflow](#)

## 数据库的表设计

数据库的设计里，我初步是想基于"年-月"设计表的，但这样的话，对于用户上传的表，就要逐一进行时间提取并标记然后分类，效率势必非常之慢，所以不可行。

为了支持用户高速批量导入，应该将用户所提交的表视为无状态的表，比较合适的是，按照快递公司进行分类导入，不过这对用户可能不太友好，毕竟有30多家公司.....

那回到之前的方案，如果是按月导入，程序不去核对月份，而按照用户对表的约定进行数据库分类，这样是否可行呢？也许这是一个不错的方案，但这对用户的要求太高了，目前用户导出的erp表的表名与内容还有对不上的，比如ljx导的表名是12月14日的表，但里面的内容其实是11月份的，这让程序或者用户进行表主体月份划定，就存在较大的误判可能。

如果什么都不做呢？直接塞表，不去计算月份，也不按月份进行表划分，也不按快递公司进行表划分，这样的话最大的问题就是表的体积越来越大，后续匹配的速度可能就会越来越慢了。毕竟也没人可以与愿意对后续的表进行维护与调优。

综合考虑的话，那还是在数据插入时，程序计算，然后自动分类到按月划分的表里，这样插入虽然慢一些，但是好歹匹配起来会比较快，而且也不太需要考虑后续维护的问题，毕竟每个月的数据量再怎么样也是可以接受的，后续匹配的效率也可以得到保障。

## [DEPRECIATED] Module Design

- Import Panel
- Analysis Panel
- ReviselnAnalysis Panel
- Upload Panel
- Comparison Panel
- Database Panel
- Feedback Panel

## 心得

Interface helps do right!

```
const onValidateError: IOnValidateError = (e1: IpcMainEvent, err: IParseError) => {
  reply(
    e1,
    channel: RequestParseFile,
    msg: makeMsgFromMain({ error: err.msg.message, content: err, level: MsgLevel.warn })
  );
};

const onData: IOnData = () => {
  // do nothing
};

const onFinish: IOnFinish = (e1: IpcMainEvent) => {
  console.log(`reading file, name: ${fp}`);
};
```

## Interface 和 Object 之间的关系

如图，在我花了漫长的时间终于设计出一个目前接口比较良好的`handleParseFile`的函数之后，为了暴露给前端，我需要写一个接口。

```
interface IHandleParseFile {
  fp: string;
  withHeader: boolean;

  onInnerError?: (e: string) => void;
  onError?: (e: IParseError) => void;
  onData?: (e: IParseData) => void;
  onFinish?: () => void;
}

export const handleParseFile = async (
  e,
  fp: string,
  withHeader: boolean = false,
  onInnerError?: (e: string) => void,
  onError?: (e: IParseError) => void,
  onData?: (e: IParseData) => void,
  onFinish?: () => void
): Promise<void | CsvParserStream<Row<any>, Row<any>>> => {
  console.log(`reading file, name: ${fp}`);
};
```

这个时候，问题来了，我当然可以直接用`ReturnType`自动追踪这个函数的参数与返回类型.....

## OH NO!

写着写着我发现我想错了，我虽然可以从函数自动得到它的返回类型，但我貌似是没法得到这个函数的参数类型的吧（待确定）！

你看，写笔记还是有帮助的，不写还不知道自己的认知其实是错的呢！

这就先去查一查，是否可以将接口自动同步于某个函数，或者说，是否可以从函数生成接口。

can ts generate interface from a function?

About 5,470,000,000 results (0.57 seconds)

<https://www.typescriptlang.org/docs/interfaces.html>

**Handbook - Interfaces - TypeScript**

Function Types — Here, we show how you can create a variable of a **function** type and assign it a **function** value of the same type. ts. let mySearch : ...

[Our First Interface](#) · [Excess Property Checks](#) · [Function Types](#) · [Indexable Types](#)

You visited this page on 12/30/21.

**People also ask :**

- Can a function implement an interface TypeScript?
- Does TypeScript support interfaces?
- Can you implement a function in an interface?
- When should I use interface TypeScript?

Feedback

<https://www.tutorialsteacher.com/typescript-interface.html>

**TypeScript Interfaces - TutorialsTeacher**

An **interface** is defined with the keyword **interface** and it can include properties and **method** declarations using a **function** or an **arrow function**.

<https://blog.logrocket.com/interfaces-in-typescript-what-are-they-and-how-do-we-use-them/>

**Interfaces in TypeScript: What are they and how do we use them**

19 Apr 2019 — Also, the **buildTeslaModelS** function doesn't know that it would have to deal ... A class or function can implement an interface to define the

确实不行！都没有相关问题！

老老实实写接口吧！（原以为是先有鸡再有蛋的问题呢！结果ts直接把“蛋生鸡”路给封了！）

```
import { checkCsvEncoding } from '../parse/utils';

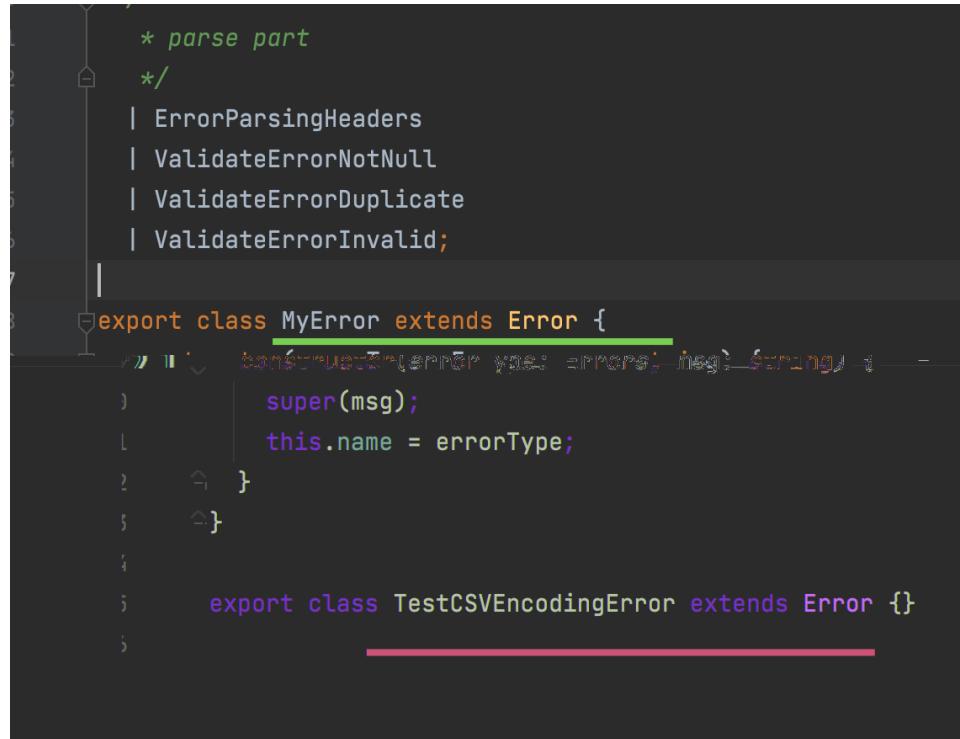
/**
 * @param e - event
 * @param fp - file path
 * @param fso - fast csv object
 * @param bo - boolean
 * @param withHeader - electron event
 * @param fp - file path
 * @param e - event
 * @param constraint - constraint
 * @param e - event
 * @param e - event
 */
const handleParseFile: (e: any, fp: string, withHeader?: boolean, onInnerError?: ((e: any, msg: string) => void) | undefined, onError?: ((e: any, err: IParseError) => void) | undefined, onData?: ((e: any, data: IParseData) => void) | undefined, onFinish?: ((e: any) => void) | undefined) =>
  void = async (e, fp, withHeader, constraint, onInnerError, onError, onData, onFinish) => {
    if (withHeader) {
      const header = await checkCsvEncoding(fp);
      if (!header) {
        return;
      }
    }
    const stream = fs.createReadStream(fp);
    const parser = fast_csv.createParser(header);
    parser.on('data', (data) => {
      if (constraint) {
        const constraintResult = constraint(data);
        if (!constraintResult) {
          return;
        }
      }
      onData?.(data);
    });
    parser.on('error', (err) => {
      onError?.(err);
    });
    parser.on('end', () => {
      onFinish?.();
    });
    stream.on('error', (err) => {
      onInnerError?.(err);
    });
    stream.on('close', () => {
      onFinish?.();
    });
  };
}

export const handleParseFile = async (e, fp, withHeader, onInnerError, onError, onData, onFinish) => {
  const header = await checkCsvEncoding(fp);
  if (!header) {
    return;
  }
  const stream = fs.createReadStream(fp);
  const parser = fast_csv.createParser(header);
  parser.on('data', (data) => {
    if (onConstraint) {
      const constraintResult = onConstraint(data);
      if (!constraintResult) {
        return;
      }
    }
    onData?.(data);
  });
  parser.on('error', (err) => {
    onError?.(err);
  });
  parser.on('end', () => {
    onFinish?.();
  });
  stream.on('error', (err) => {
    onInnerError?.(err);
  });
  stream.on('close', () => {
    onFinish?.();
  });
}
```

## Error类的继承设计

很有意思的一点，就是在整合优化代码时，发现整合到了Error类的冗余定义。

再仔细想想eslint对每个文件只能有一个class导出的约定，还挺有意思的，把类拆开，代码变得更加好组织了。

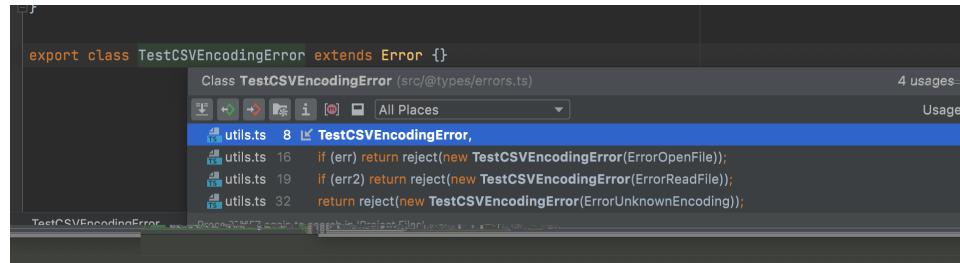


```
* parse part
*/
| ErrorParsingHeaders
| ValidateErrorNotNull
| ValidateErrorDuplicate
| ValidateErrorInvalid;

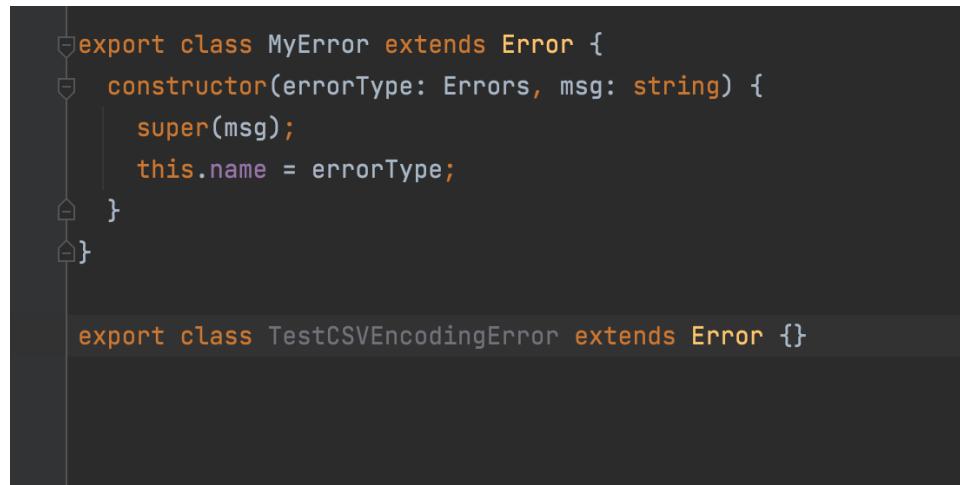
export class MyError extends Error {
    constructor(errorType: Errors, msg: string) {
        super(msg);
        this.name = errorType;
    }
}

export class TestCSVEncodingError extends Error {}
```

所以接下来就是把所有用这个类的代码都消掉了。



OK，很快就消除完了，毕竟我后续定义的MyError类是这个TestCsv...类的超集。



```
export class MyError extends Error {
    constructor(errorType: Errors, msg: string) {
        super(msg);
        this.name = errorType;
    }
}

export class TestCSVEncodingError extends Error {}
```

[ARCHIVE] deploy script

This section is done in the primary era of this project, and the solution is based on web, which is proved to be not suitable later (but valuable for reuse).

Now the project is fully based on **Electron + React + TypeScript + Sqlite3** (a minor local database).

### deploy backend (python, fastapi)

suppose you have deployed project onto the server under directory **hjxh\_express\_match/backend/**, the backend structure may be:

```
.  
├── config.py  
├── const.py  
├── db.py  
├── hash.py  
├── log.py  
└── logs  
├── main.py  
├── requirements.txt  
├── static  
└── templates
```

the **requirements.txt** includes all the dependencies need by the python interpreter of this project.

We should first create a new environment on this server, for pure use of this project, i.e:

1. install the python on the server, the version of which would better correspond with the one of the local in case of unexpected error caused by version difference
2. use **virtualenv** to create an env based on this python version named **venv\_py** under this working directory
3. activate this env
4. use **pip** to install the **requirements.txt**
5. run!

```
PY_VERSION=python3.8

# install the target python version based on its version number
# if you don't use these two lines, then you would suffer from `wget
blablabla...` when you checked what the hell the python repo url is
sudo apt install software-properties-common -y
sudo add-apt-repository ppa:deadsnakes/ppa
sudo apt install ${PY_VERSION}

# use `virtualenv` to create and activate a new python env fast~
sudo apt install virtualenv
virtualenv -p ${PY_VERSION} venv_py
source venv_py/bin/activate
```

```
# install all the requirements
# if you need to dump all the requirements of a python project used, you
can use `pip freeze > requirements.txt` so that a file named of
`requirements.txt` would be generated under the current directory
pip install -r requirements.txt

# run our backend of `fastapi`
python main.py
```

deploy frontend (node, react)

Install nodejs: version > xx

Install npm

Install yarn

1. install nodejs
2. install npm
3. install yarn: `sudo npm install -g yarn`

reference:

- [mermaid](#) - Markdownish syntax for generating flowcharts, sequence diagrams, class diagrams, gantt charts and git graphs.

## [ARCHIVE] 表约定【重要】

该章节的全部内容均已经更新，并且独立成单独的文件，方便用于前端展示，已不在总文档内更新，具体可见：

- [upload\\_base](#)
- [upload\\_erp](#)
- [upload\\_trd](#)

## 术语定义

- erp表：erp数据由erp系统导出，以`.csv`文件形式，内部只会有一张表，该文件即该表即erp表即erp的明细表
- trd表：trd数据由各物流公司给出，以`.xlsx|.xls`文件形式，内部可能会有多张表，其中包含物流明细信息的表叫明细表，该明细表即trd表

## 通用表约定

对于任意一张明细表，尽管即使表的内容存在一定的不规范也不影响解析，但总体来说，表的质量也有优劣之分，以下将给出详细的参照，在系统读表的过程中如果出现问题可以对照该参照，一定程度上可以找到问题的所在。

1. 明细表的第一行，必须是列字段，不得出现第一行为 、为非相关信息等，否则将出错
2. 明细表的数据主体必须是一个完整的矩形，即列与列之间不得 列、行与行之间不得有 行，否则将导致数据解析丢失
3. 明细表的数据主体之后，允许有多余的其他汇总信息，但是，这可能导致最终解析错误，建议不要加 汇总信息，或者至少隔一行。如果您们使用excle自带的统计汇总功能，请剪切最后一行然后粘贴到第 N+2行或以后；如果直接用右键插入一行 行，实际上程序依旧会把它们识别成一起，最终导致错误。
4. 基于以上格式规范的表，是一张较为完美的表。

## erp表约定

1. erp表必须以.csv格式导出，并且加上erp\_前缀，否则无法解析
2. erp表中必须包含以下列字段：
  1. 物流单号 --> \_id
  2. 收货地区 --> area\_erp
  3. 实际重量 --> weight\_erp
  4. 发货时间 --> time\_erp
  5. 物流公司 --> logistics\_erp

## trd表约定

1. trd表必须以.xlsx | .xls格式导出，并且加上trd\_前缀，否则无法解析
2. trd表中必须有一张明细表，该明细表必须有

3. 快递单号还发现有"=xxxx"格式存储的，可能是用了公式啥的，尽管已经设计了算法还原，但是尽量避免这样的问题产生
  4. 否则该行报错
2. 重量 (weight\_erp, weight\_trd)
    1. 所有重量不得为空
    2. 所有重量必须是一个浮点数（小数或整数）
    3. 所有重量必须大于0
    4. 否则该行报错
  3. 省份 (area\_erp, area\_trd)
    1. 所有省份不得为空
    2. 所有省份必须以合法的34个省市自治区的前缀开头
    3. 否则该行报错
  4. 发货时间 (time\_erp, time\_trd)
    1. 所有时不得为空
    2. 所有时间必须包含"YYYY.MM.DD"或"MM.DD.YYYY"字符串样式，其中"."可以是"-"或者"/"
    3. 实际解析过程中发现还有以1900年起的天数作为日期列的，即4w+的一个数，尽管我已经写了一个逆转算法，但是尽量避免这样的问题产生，这里面涉及到excel中短日期的显示问题，具体可以查一下相关资料
    4. 否则该行报错