

# Potato Audit

## *Get Started*

**In Command**

**cd to target directory and run *npm start***

```
C:\Users\shiyaom>cd Desktop/Shiyao  
  
C:\Users\shiyaom\Desktop\Shiyao>cd potato-audit  
  
C:\Users\shiyaom\Desktop\Shiyao\potato-audit>npm start
```

**cd to src and run *python app.py***

```
C:\Users\shiyaom>cd Desktop/Shiyao/potato-audit  
  
C:\Users\shiyaom\Desktop\Shiyao\potato-audit>cd src  
  
C:\Users\shiyaom\Desktop\Shiyao\potato-audit\src>python app.py
```

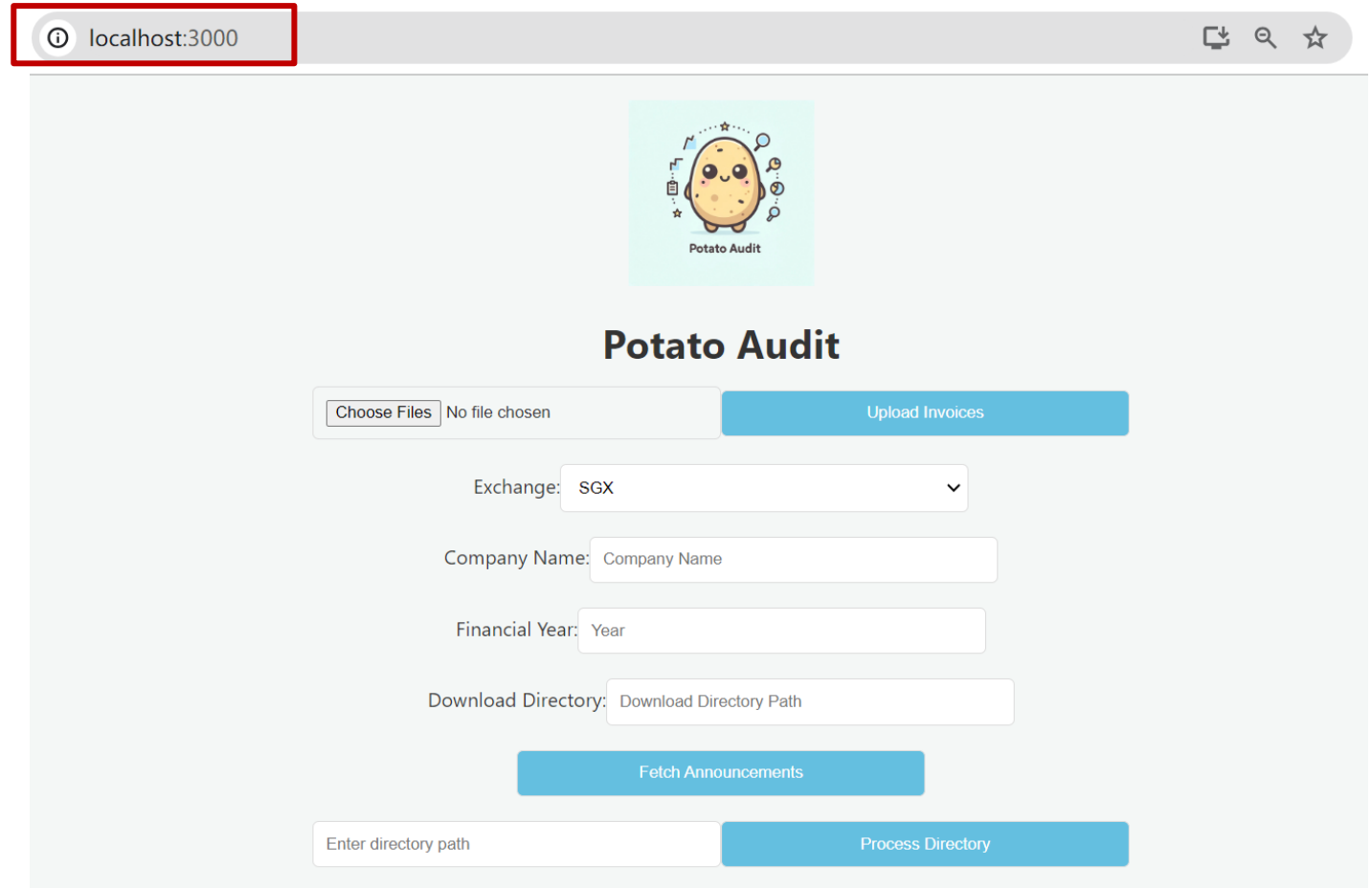
**To Do:**

**Build package for the app so it could be started with a click**

# Potato Audit

## *Get Started*

**In Browser (Chrome)**  
***Web UI is running on localhost 3000***



The screenshot shows a web browser window with the address bar displaying "localhost:3000". The page features a logo of a potato character with the text "Potato Audit" below it. The main heading "Potato Audit" is centered. Below the heading, there is a file upload section with a "Choose Files" button and a "No file chosen" message, followed by an "Upload Invoices" button. Below this, there are input fields for "Exchange:" (set to "SGX"), "Company Name:" (with placeholder "Company Name"), "Financial Year:" (with placeholder "Year"), and "Download Directory:" (with placeholder "Download Directory Path"). At the bottom, there is a "Fetch Announcements" button and a section with an "Enter directory path" input field and a "Process Directory" button.

localhost:3000

Potato Audit

Choose Files No file chosen Upload Invoices

Exchange: SGX

Company Name: Company Name

Financial Year: Year

Download Directory: Download Directory Path

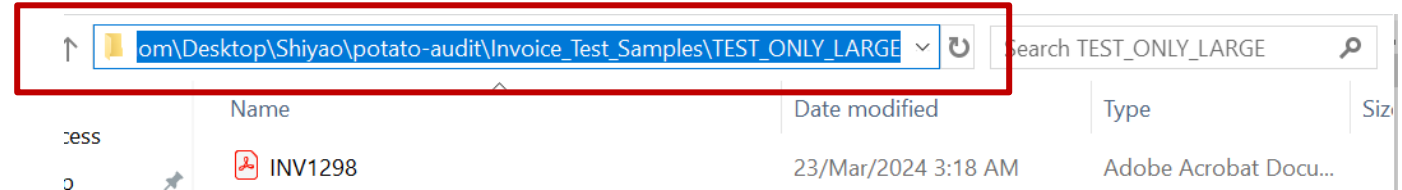
Fetch Announcements

Enter directory path Process Directory

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

Go to invoice folder



Copy and paste directory path  
to input field

localhost:3000

Potato Audit

Choose Files No file chosen Upload Invoices

Exchange: SGX

Company Name: Company Name

Financial Year: Year

Download Directory: Download Directory Path

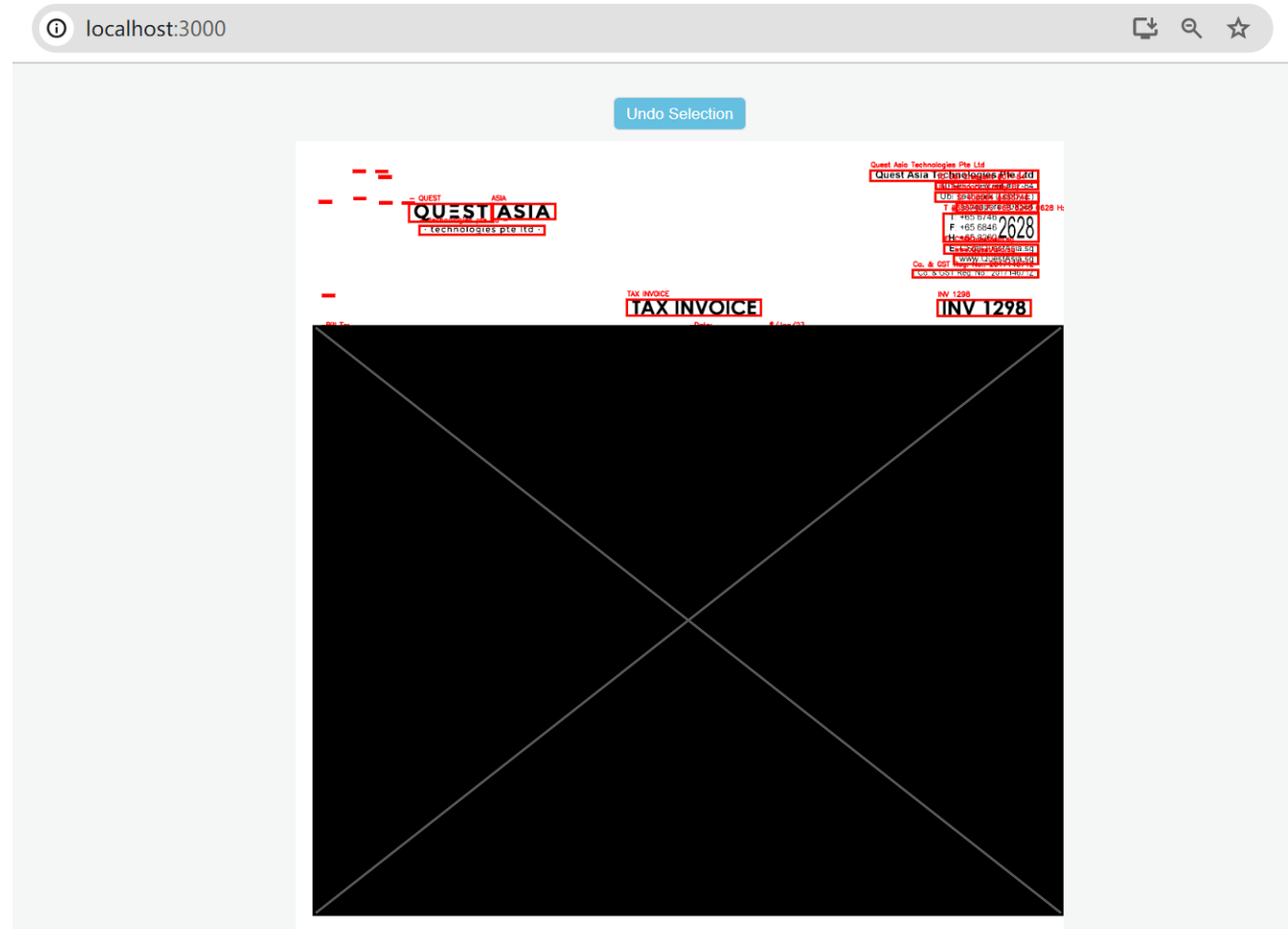
Fetch Announcements

Process Directory

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

Scroll down to find invoice  
image with annotated *bounding*  
*boxes*



# Potato Audit

## To Extract Data from Invoice/Structured Documents

Enlarge the page if needed  
Click on boxes to select *Field Label* (A) and *Value* (B) pairs that we need

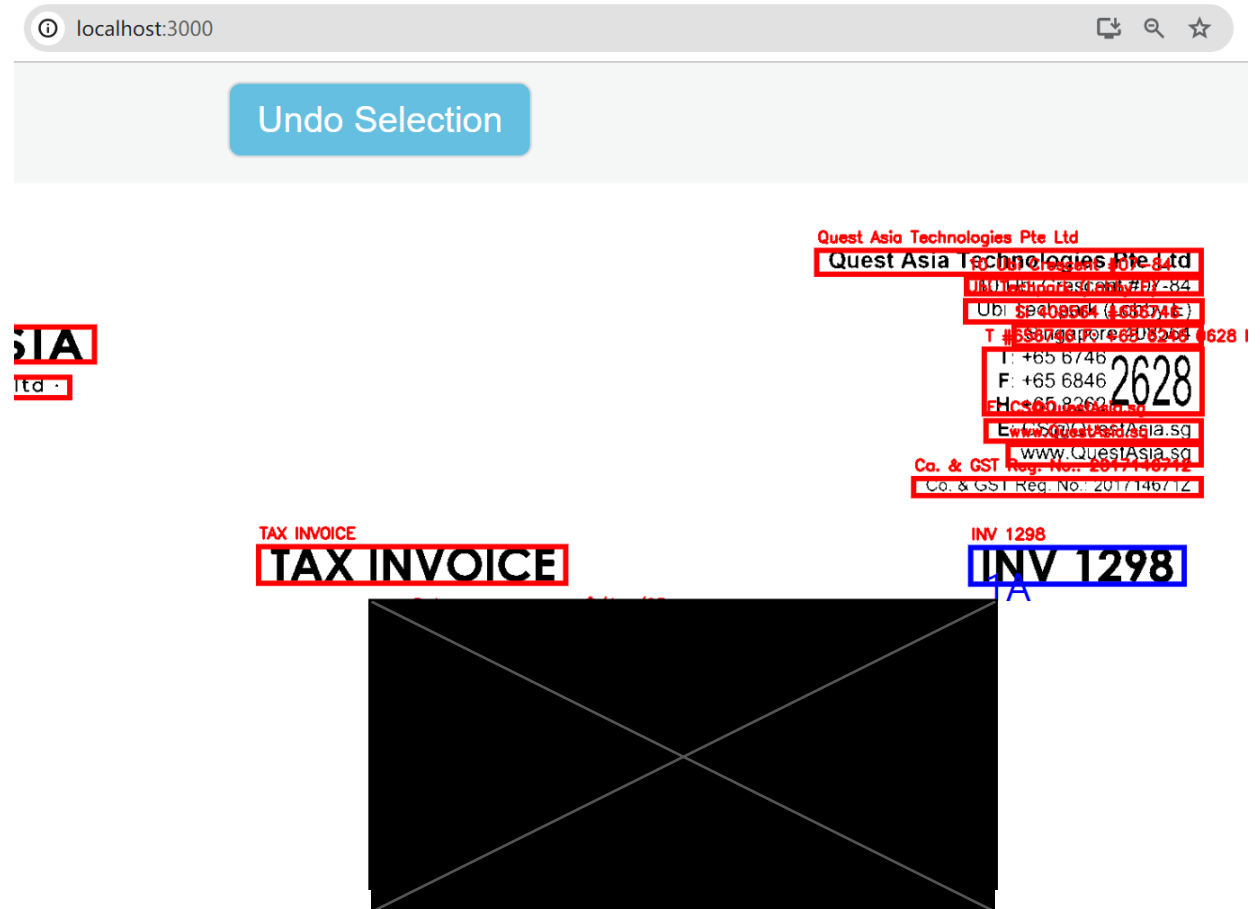
If a *Field Label* is also a *Value* itself, click on it twice

Example:

Field Label: INV

Value: 1298

They are in the same box, so click on it twice



# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

Once a pair is selected, enter one or more *Field Label* variations, if any, separated by comma

**Example:**

**Field Label: Bill To**

**Possible variations: Sold To**

The screenshot displays a sample invoice titled "TAX INVOICE" with various fields highlighted by red bounding boxes. These fields include "To:", "Date:", "Invoice No.", "Signed Quote dated", "Your PO #", "GST Reg No.", "INV 1298", "TA ID", "Date: 6/Jan/23", "Invoice No: 1298", "Signed Quote dated 6/12/2021", and "Your PO #".

Below the invoice, there is a form for entering field label variations. The form has a text input area with the placeholder text "Field label variations, separated by comma - e.g., Invoice, Bill, Receipt". Below the input area are two buttons: "Save" and "No need".

## ***To Extract Data from Invoice/Structured Documents***

TAX INVOICE

**TAX INVOICE**

INV 1298  
1A 1B

To:

**WOMEN'S CLINIC FOR WOMEN PTE LTD**

**WOMEN'S CLINIC FOR WOMEN PTE LTD**

**Women's Medical Centre #05-04**

**88V, Brasen Road**

**Singapore 307677**

**Attn: Dr A L Lim**

Date:

**6/Jan/23**

**6/Jan/23**

**6/Jan/23**

**6/Jan/23**

**Your PO #**

**Signed Quote dated 6/12/2021**

**Signed Quote dated 6/12/2021**

Sold To:

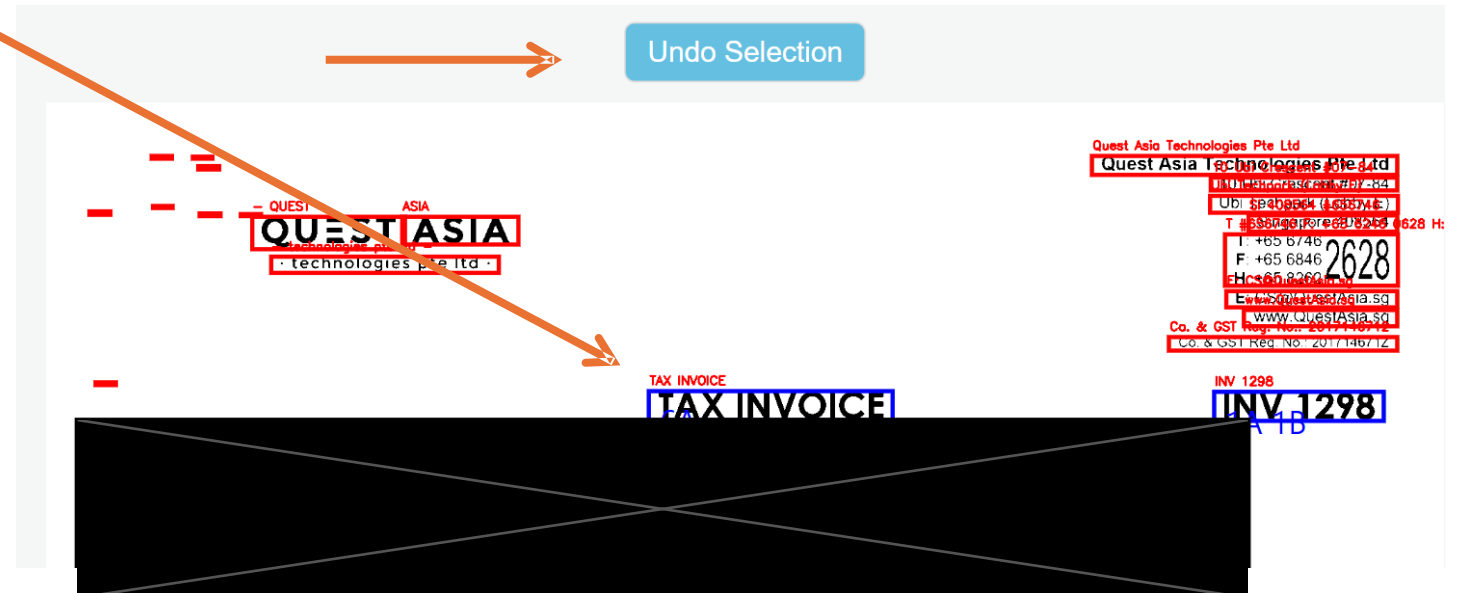
Save

No need

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

If mistakenly selected a box,  
simply click *Undo Selection* to  
revert the change

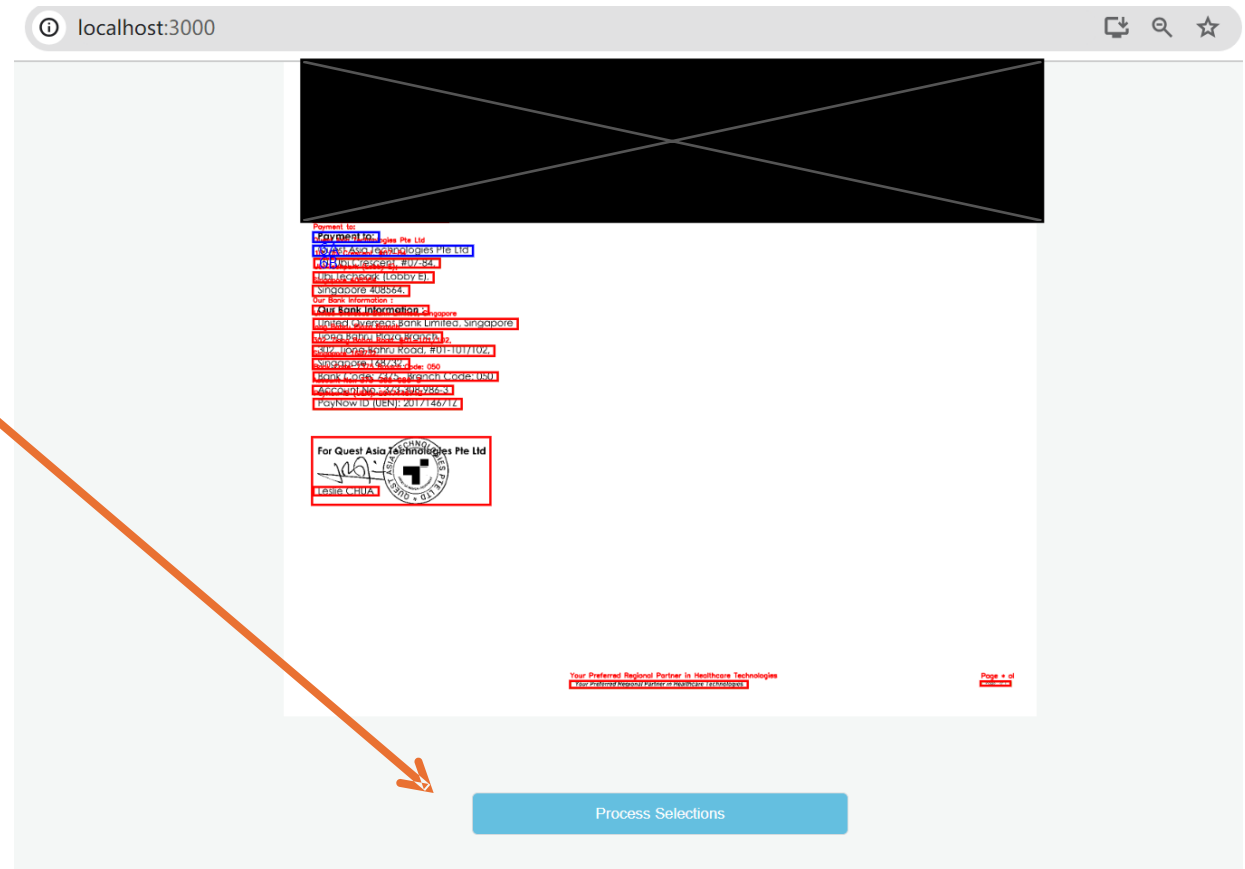




# Potato Audit

## To Extract Data from Invoice/Structured Documents

Upon selecting pairs needed,  
click *Process Selection*

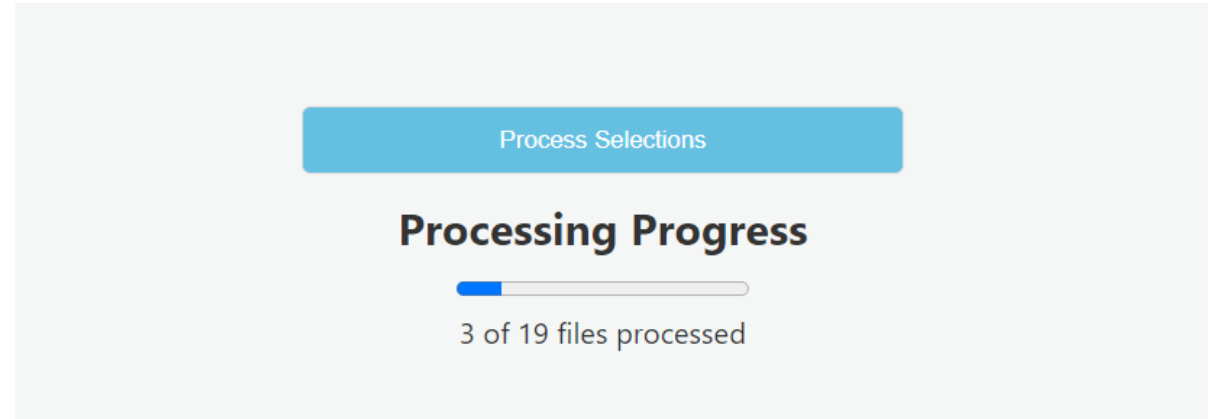


# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

It is processing all the files in the specified directory, assuming they have similar format

Progress bar appears to track the status



# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

Once completed, a table appears below with sorting and search functions for preliminary inspection

Due to formatting issues, items may not be sorted correctly here


If there is missing or incorrect entries, copy the *File URL* in the last column to File Explorer to check the original file



Processing Progress					
<div><div></div></div>					
19 of 19 files processed					
Search: 19 records...					
Bill To:	Date:	GRAND TOTAL \$	INV	Payment to:	Sub-Total \$ <input checked="" type="checkbox"/>
File URL					
					<a href="#">C:\Users\jason\</a>

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

For better post-processing, an Excel file with extracted data is saved to the same invoice directory named 'processed\_invoices'

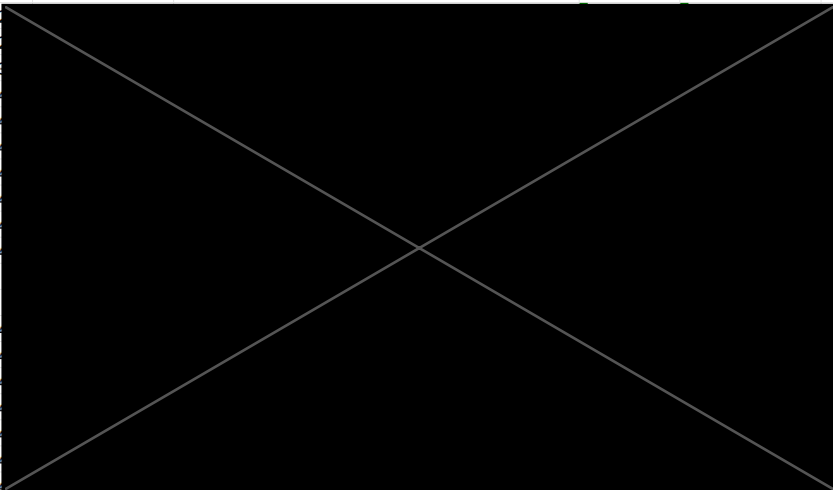


 processed_info	3/29/2024 1:36 AM	JSON 源文件	8 KB
 processed_invoices	3/29/2024 1:36 AM	Microsoft Excel 工...	7 KB

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

Open the Excel file to further  
clean and process extracted data



	A	B	C	D	E	F	G
1	INV	Date:	Bill To:	Sub-Total \$	GRAND TOTAL \$	Payment to:	File URL
2	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
3	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
4	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
5	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
6	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
7	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
8	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
9	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
10	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
11	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
12						uest Asia Technologies Pte Ltd	Click Here to Open File
13						uest Asia Technologies Pte Ltd	Click Here to Open File
14	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
15	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
16	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
17	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
18	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
19	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
20	INV 1					uest Asia Technologies Pte Ltd	Click Here to Open File
21							

# Potato Audit

## *To Extract Data from Invoice/Structured Documents*

If there is any missing or incorrect data, simply click on the cell of the last column in given row, which directly pulls up the original file

10	INV 1435		Click Here to Open File
11	INV 1436		Click Here to Open File
12			Click Here to Open File
13			Click Here to Open File
14	INV 1438		Click Here to Open File
15	INV 1438		Click Here to Open File
16	INV 1439		Click Here to Open File

Look for the information in the original file

TEL: 65 6202 1111  
E: [CS@QuestAsia.sg](mailto:CS@QuestAsia.sg)  
[www.QuestAsia.sg](http://www.QuestAsia.sg)

Co. & GST Reg No.: 201714671Z

**TAX INVOICE**

**INV 1437**

Date:

Our ref:

Your PO #

Ref :

# Potato Audit

## *Performance Metrics*

Training dataset of 175 sales invoices

							Asia Technologies Pte Ltd	Click Here to Open File
177	Invoice Number	Buyer	Date	Sub-Total	Grand Total	Seller		
178	158		175	170	140	173	173	
179	90%		100%	97%	80%	99%	99%	
180			Adjusted for blank sub-totals ->	173				
181				99%				

Processing time for training set ~350s

Processing time per invoice ~ <2s

To Do:

Test with invoices of other formats/languages

# Potato Audit

## *Potential Use Cases with Invoice Data*

**Vouching and Reconciliation**

**100% Coverage of Samples**

**Visualise spending or sales patterns - check for anomaly (unusual volume/amount/date of transaction, duplicate payments, round amount invoices, or payments to unfamiliar vendors)**



# Potato Audit

## *To Fetch Announcements for Listed Co.*


Select the exchange of the company (supports SGX/ASX; may integrate more exchanges if needed)

Enter the name/ID as recorded on the exchange (may consider doing a fuzzy match in the future)

Enter the financial year required

Specify a directory for the files to be downloaded to

Hit *Fetch Announcements*



### Potato Audit

No file chosen

Exchange:

Company Name:

Financial Year:

Download Directory:

# Potato Audit













## *To Fetch Announcements for Listed Co.*

Folder of  
<name of company>\_<financial  
year> created at specified  
directory

Auto renaming of files with  
datetime prepended for easy  
sorting

Wait for the downloads

CSV file with extracted content  
created named  
'Extracted\_Announcements'

OEL_(HOLDINGS)_LIMITED_FY20243/22/2024 8:34 PM文件夹				
名称	修改日期	类型	大小	
 20240108_2254_OEL - Annt - Proposed Ca...	3/3/2024 3:26 AM	Microsoft Edge PD...	140 KB	
 20240111_0759_OEL Holdings - Circular to ...	3/3/2024 3:26 AM	Microsoft Edge PD...	931 KB	
 20240111_0759_OEL Holdings - Notice of E...	3/3/2024 3:25 AM	Microsoft Edge PD...	127 KB	
 20240111_0759_OEL Holdings - Proxy Form	3/3/2024 3:25 AM	Microsoft Edge PD...	32 KB	
 20240117_2157_OEL Holdings - Proposed ...	3/3/2024 3:25 AM	Microsoft Edge PD...	137 KB	
 20240118_1844_OEL - Ann - Response to Q...	3/3/2024 3:25 AM	Microsoft Edge PD...	140 KB	
 20240206_2350_OEL Holdings - Circular to ...	3/3/2024 3:25 AM	Microsoft Edge PD...	931 KB	
 20240206_2350_OEL Holdings - Notice of E...	3/3/2024 3:25 AM	Microsoft Edge PD...	127 KB	
 20240206_2350_OEL Holdings - Proxy Form	3/3/2024 3:25 AM	Microsoft Edge PD...	32 KB	
 20240206_2350_OEL Holdings - Results of ...	3/3/2024 3:25 AM	Microsoft Edge PD...	126 KB	
 20240219_1817_OEL Holdings - Change of ...	3/3/2024 3:25 AM	Microsoft Edge PD...	37 KB	
 Extracted_Announcements	3/3/2024 3:33 AM	Microsoft Excel 工...	53 KB	

# Potato Audit

## To Fetch Announcements for Listed Co.

Require some manual reformatting -  
could be enhanced by passing the  
content to LLM for a summary  
if needed

	A	B	C	D	E	F
1	Date	Ref	Category	Details	FS Impact/AWP	
				<p>1</p> <p>OEL (HOLDINGS) LIMITED (Incorporated in the Republic of Singapore) (Unique Entity Number: 198403368H)</p> <hr/> <p>PROPOSED CAPITALISATION</p> <hr/> <p>1. INTRODUCTION</p> <p>The Board of Directors (the "Board") of OEL (Holdings) Limited (the "Company" and together with its subsidiaries, the "Group") wishes to announce that the Company has on 8 January 2024 entered into the following:</p> <p>(a)</p> <p>a conditional capitalisation deed dated 8 January 2024 (the "ZJ Conditional Capitalisation Deed") with Dr. Zhang Jian, the Executive Chairman and a substantial shareholder of the Company, pursuant to which the Company agreed, subject to and upon the terms and conditions set out therein, to capitalise the ZJ Loan Outstanding Amount (as defined herein) via the allotment and issuance of 145,285,714 new Shares (as defined herein) (the "ZJ Capitalisation Shares") at \$50.007 per Share (the "Capitalisation Price") to Dr. Zhang Jian, credited as fully paid-up and free from any mortgage, charge, security interest, lien, pledge, assignment by way of security, equity,</p>		
2	08 Jan 2024 10:54 PM	ES13-1	OEL - Annt - Proposed Capitalisation	<p>CIRCULAR DATED 11 JANUARY 2024</p> <p>THIS CIRCULAR IS IMPORTANT AND REQUIRES YOUR IMMEDIATE ATTENTION. PLEASE READ IT CAREFULLY.</p> <p>If you are in any doubt as to the course of action you should take, you should consult your stockbroker, bank manager, solicitor, accountant or other professional adviser(s) immediately.</p> <p>If you have sold or transferred all your shares in the capital of OEL (Holdings) Limited (the "Company") held through CDP (as defined herein), you need not forward this Circular, the Notice of EGM (as defined herein) and the Proxy Form (as defined herein) to the purchaser or transferee as arrangements will be made by CDP for a separate Circular to be sent to the purchaser or transferee. If you have sold or transferred all your shares in the capital of the Company by physical share certificate(s), you should immediately forward this Circular, the</p>		

# Potato Audit

## *Performance Metrics*

**Downloading time for 50+ files (or any number of files for an entire year)  
~ 2-3min max depending on Internet speed**

**Text extraction to CSV ~ < 10s**

**Manual Reformatting ~ < 10min**

**Note:**

**Texts extracted are the same as the announcements**

**Each announcement occupies one row (may consider using VBA to reformat things)**