

Intelligent Document Extraction with OpenAI GPT & LangChain

OpenAI API | GPT | Vision Models | MongoDB | LangChain | FastAPI

Get Started



Invoice

amazon.in

Tax Invoice/Bill of Supply/Cash Memo
(Original for Recipient)

Sold By :
CLICKTECH RETAIL PRIVATE LIMITED
Survey Number 99/1, Mamidipally Village,
Shamshabad
HYDERABAD, TELANGANA, 500108
IN

PAN No:AAJCC9783E
GST Registration No:36AAJCC9783E1Z8
Dynamic QR Code:

Order Number:407-7033233-4757902
Order Date:30.11.2024

Sl. No.	Description	Unit Price	Qty	Net Amount	Tax Rate	Tax Type	Tax Amount	Total Amount
1	Corsair Katar Pro Ultra-Light Optical USB Gaming Mouse with Backlit RGB LED, 12400 DPI (Black) B08KSZV3SR (B08KSZV3SR) HSN:84716060	₹1,143.22	1	₹1,143.22	0%	CGST	₹102.80	₹1,349.00
TOTAL:				₹205.78	9%	SGST	₹102.80	₹1,349.00

Amount in Words:
One Thousand Three Hundred Forty-nine only

For CLICKTECH RETAIL PRIVATE LIMITED:

Authorized Signatory

Whether tax is payable under reverse charge - No

*ASSPL-Amazon Seller Services Pvt. Ltd., ARIPL-Amazon Retail India Pvt. Ltd. (only where Amazon Retail India Pvt. Ltd. fulfillment center is co-located)
Customers desirous of availing input GST credit are requested to create a business account and purchase on Amazon.in/business from Business eligible offers
Please note that this invoice is not a demand for payment
Page 1 of 1

Receipt

ACE
The helpful place.

Thank You for Shopping at
Ace Hardware of Navarre
8188 Navarre Pkwy, Navarre, FL 32566
Phone: (850) 939-9800

11/14/2024, 5:46:31 PM

Reg: 2 ACE Rewards # : 1928777
Customer: J Manchester: Since 2011

# 2489540 DeWalt 20V MAX Cordless Brushed 2 Tool Kit	\$219.99
# 2029735 DeWalt Maxfit Screwdriver Bit Set 30 pc	\$19.98
# CLDBVG COKE 20oz	\$1.99
Subtotal	\$241.97
Tax	\$16.93
Total	\$258.90

Card number **** * 3081
Card type Debit
Card entry Chip
Date/time 11/14/2024 5:46 PM
Reference # 62845289260246240685C
Status APPROVED

THANK YOU
HAVE A NICE DAY



Business Card



Government ID Cards



Bank Documents

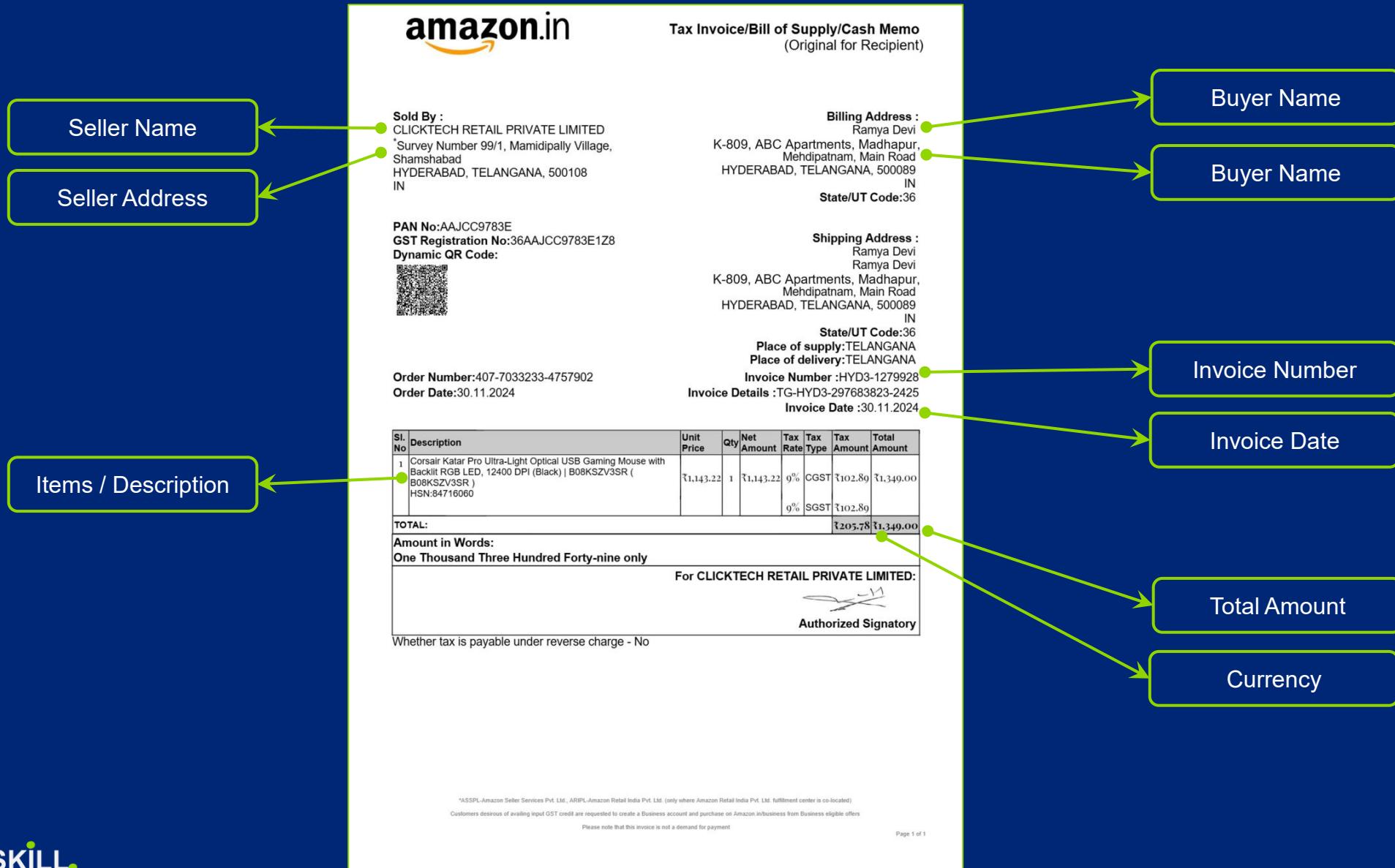
Companies handles

- Thousands of invoices
- Thousands of Receipts
- Many PDF files
- Many Scanned documents

Can Computer understand documents like humans ?

No

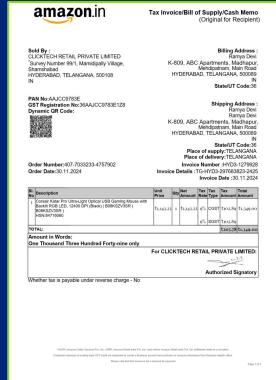
Human Brain Naturally Understand documents



What you will learn



FastAPI



Next



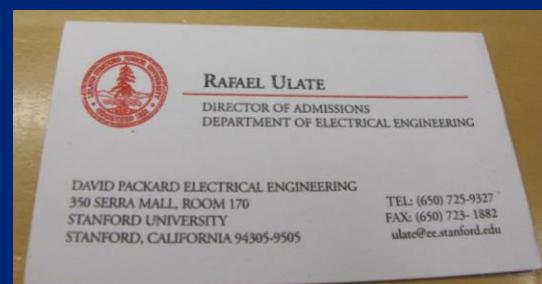
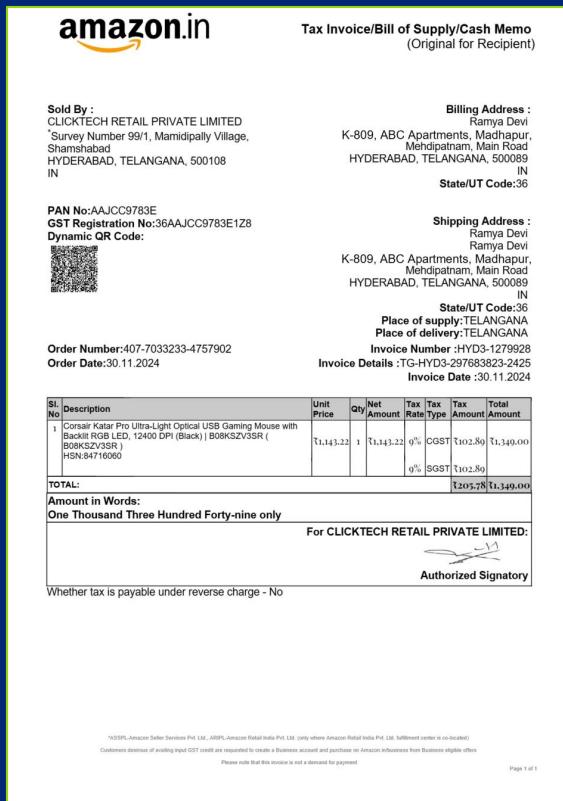
Course Plan



Intelligent Document Extraction with OpenAI GPT & LangChain

OpenAI API | GPT | Vision Models | MongoDB | LangChain | FastAPI

Course Plan



Documents

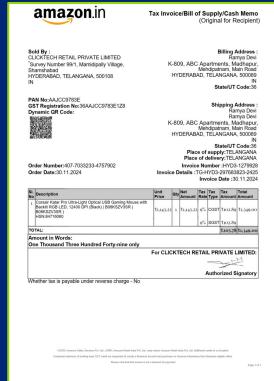
Most common documents in our life

Business & Commercial	Personal Identify Documents (India)	Customer Details
Invoice	Aadhaar Card	Electrical/Water/Telephone Bill
Shipping Invoice	PAN card	Gas Bill
Bill of Lading	Voter ID	Property Tax
Purchase Order	Driving License	Rental Agreement
Proforma Invoice	Ration Card	Agreement of Sale
Receipt of payment	Bank Passbook	Pay slips
Debit/Credit Note	Employee Photo ID card	
Worker Order	Government Issue Card	
Contract or Agreement	Health Insurance Card	
Business card		

Common Framework



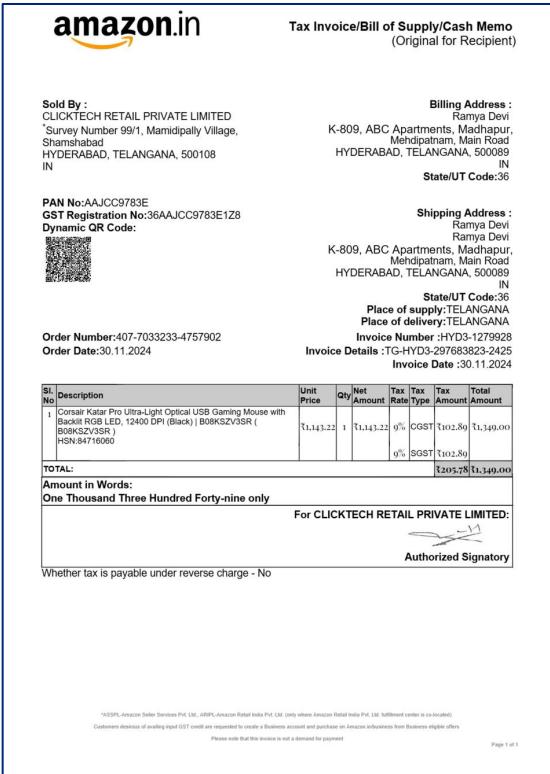
FastAPI



Documents

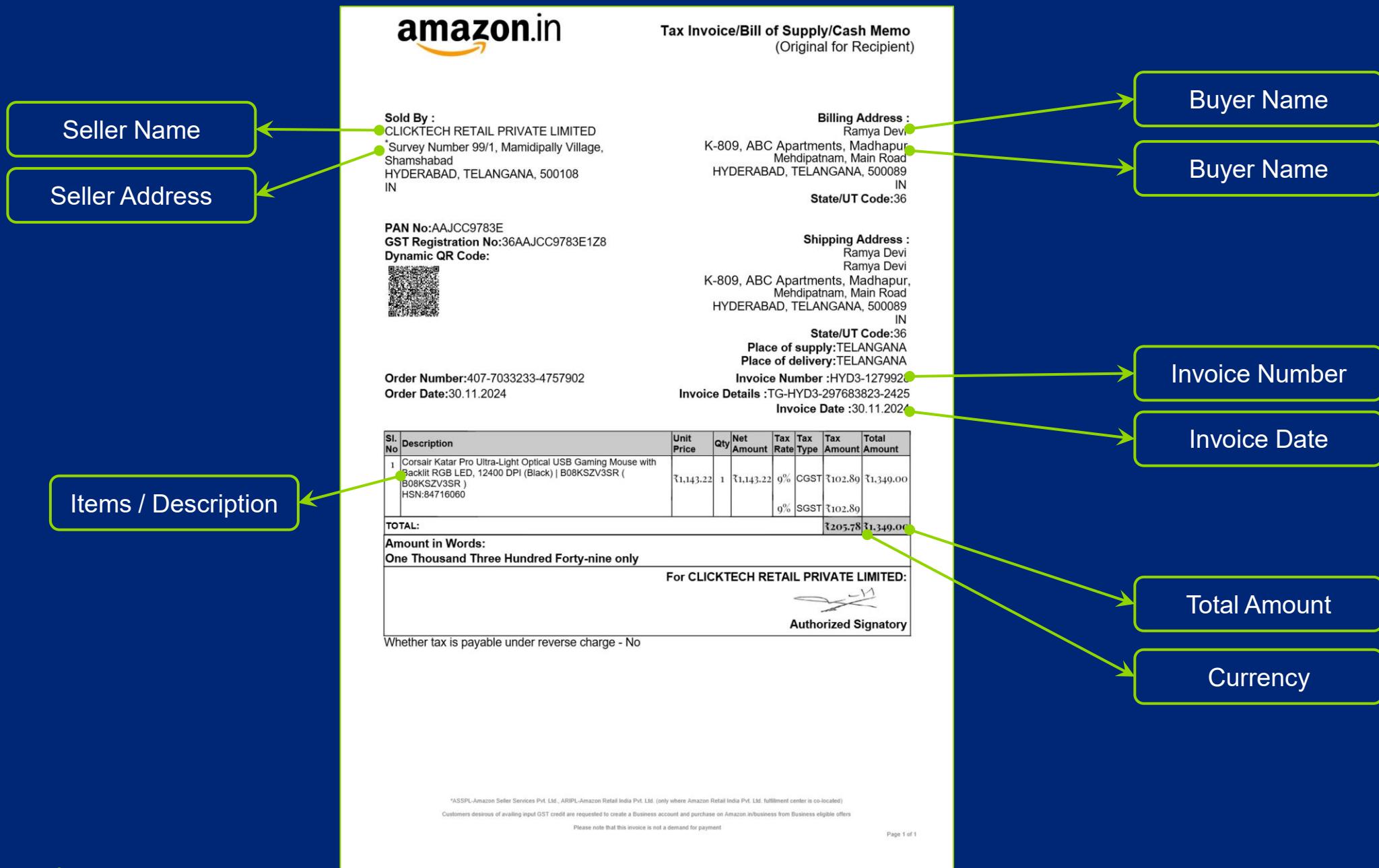
In this course we will focus on two documents

Invoice



Business Card







Tax Invoice/Bill of Supply/Cash Memo
(Original for Recipient)

Sold By :
CLICKTECH RETAIL PRIVATE LIMITED
Survey Number 99/1, Mamidipally Village,
Shamshabad
HYDERABAD, TELANGANA, 500108
IN

PAN No:AAJCC9783E
GST Registration No:36AAJCC9783E1Z8
Dynamic QR Code:



Order Number:407-7033233-4757902
Order Date:30.11.2024

Billing Address :
Ramya Devi
K-809, ABC Apartments, Madhapur,
Mehdipatnam, Main Road
HYDERABAD, TELANGANA, 500089
IN
State/UT Code:36

Shipping Address :
Ramya Devi
Ramya Devi
K-809, ABC Apartments, Madhapur,
Mehdipatnam, Main Road
HYDERABAD, TELANGANA, 500089
IN
State/UT Code:36

Place of supply:TELANGANA
Place of delivery:TELANGANA
Invoice Number :HYD3-1279928
Invoice Details :TG-HYD3-297683823-2425
Invoice Date :30.11.2024

Sl. No	Description	Unit Price	Qty	Net Amount	Tax Rate	Tax Type	Tax Amount	Total Amount
1	Corsair Katar Pro Ultra-Light Optical USB Gaming Mouse with Backlit RGB LED, 12400 DPI (Black) B08KSZV3SR (B08KSZV3SR) HSN:84716060	₹1,143.22	1	₹1,143.22	9%	CGST	₹102.89	₹1,349.00
TOTAL:								₹205.78 ₹1,349.00

Amount in Words:
One Thousand Three Hundred Forty-nine only

For CLICKTECH RETAIL PRIVATE LIMITED:

Authorized Signatory

Whether tax is payable under reverse charge - No

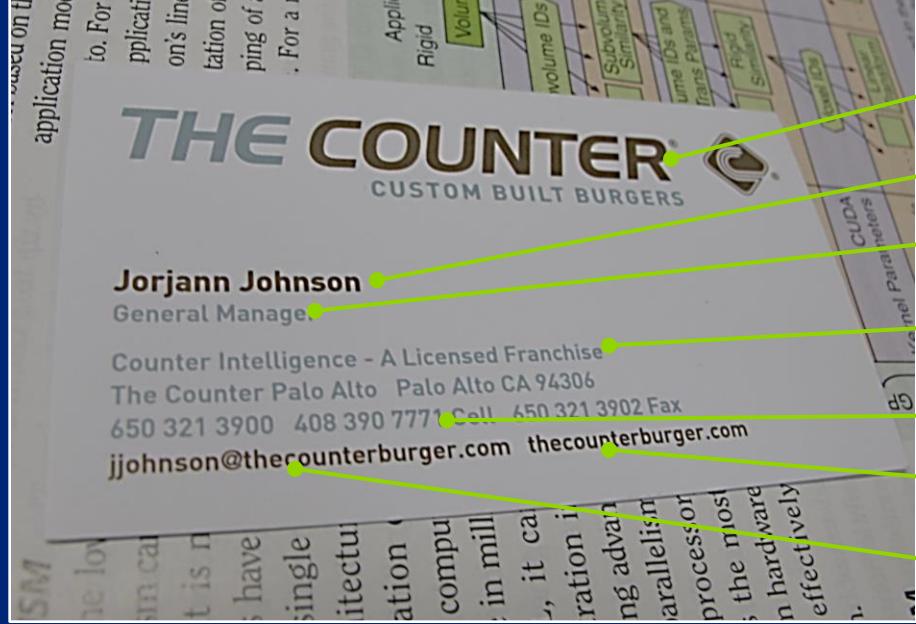
*ASSPL-Amazon Seller Services Pvt. Ltd., ARIPL-Amazon Retail India Pvt. Ltd. (only where Amazon Retail India Pvt. Ltd. fulfillment center is co-located)

Customers desirous of availing input GST credit are requested to create a Business account and purchase on Amazon inBusiness from Business eligible offers

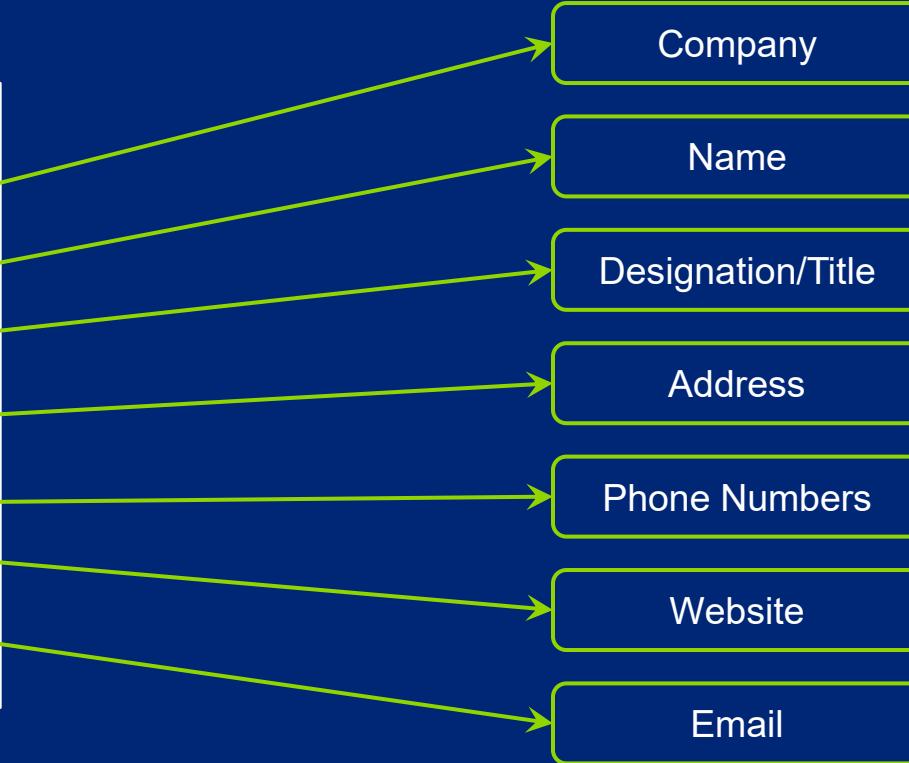
Please note that this invoice is not a demand for payment

Output from OpenAI GPT Model

```
{  
  "invoice_number": "HYD3-1279928",  
  "invoice_date": "30.11.2024"  
  "seller_name": "CLICKTECH RETAIL PRIVATE LIMITED"  
  "buyer_name": "Ramya Devi",  
  "seller_address": [  
    "SURVEY NUMBER 99/1, Mamidipally village,  
    Shamshabad, HYDERABAD, TELANGANA,  
    500108, IN"  
  ],  
  "buyer_address": [  
    "K-809, ABC Apartments, Madhapur,  
    Mehdipatnam, Main Road, HYDERABAD,  
    TELANGANA, 500089, IN"  
  ],  
  "total_amount": "1349.00",  
  "currency": "INR",  
  "items": [  
    {  
      "item_number": "1",  
      "description": [  
        "Corsair Katar Pro Ultra-Light Optical USB  
        Gaming Mouse with Backlit RGB LED, 12400 DPI  
        (black)"  
      ],  
      "quantity": "1",  
      "item_amount": "1349.00"  
    }  
  ]  
}
```



Business Card





Business Card

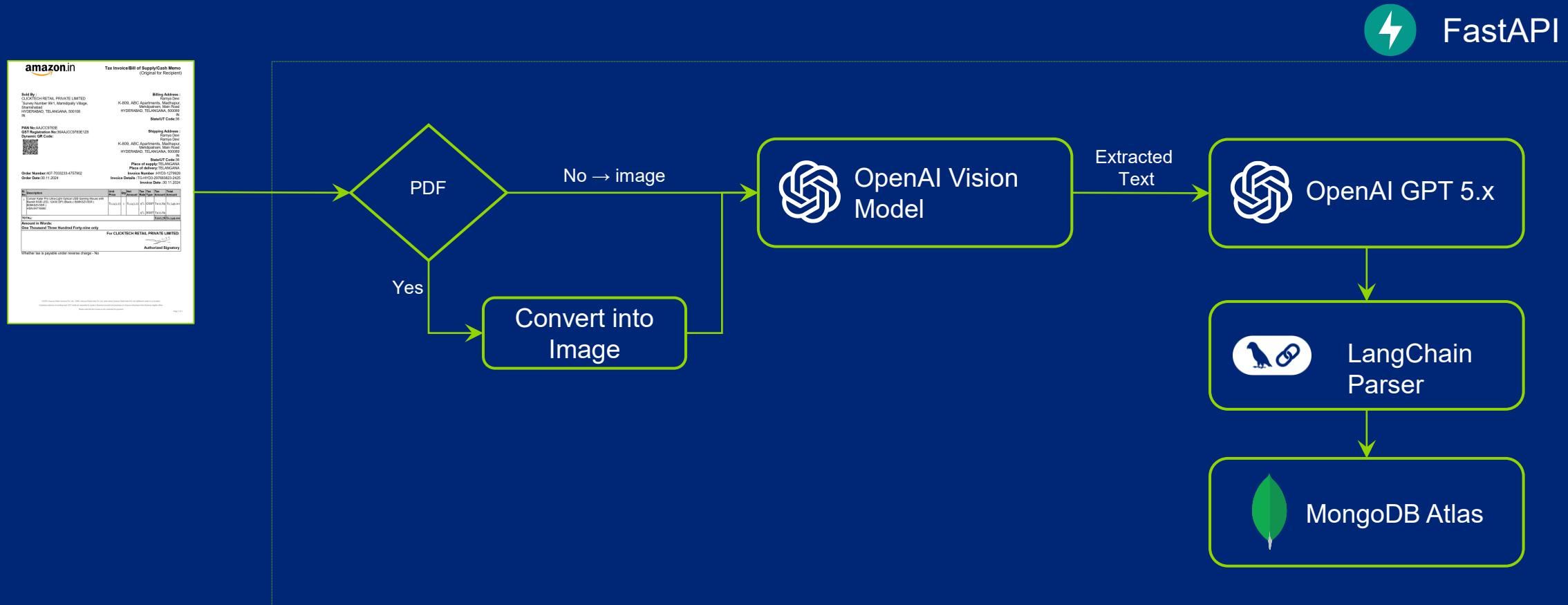


```
{  
  "name": "Jorjann Johnson",  
  "title": "General Manager",  
  "company": "The Counter",  
  "phone_numbers": [  
    "650-321-3900",  
    "408-390-7771",  
    "650-321-3902"  
  ],  
  "emails": [  
    "jjohnson@thecounterburger.com"  
  ],  
  "website": "thecounterburger.com",  
  "address": [  
    "The Counter, Palo Alto, CA 94306"  
  ],  
}
```

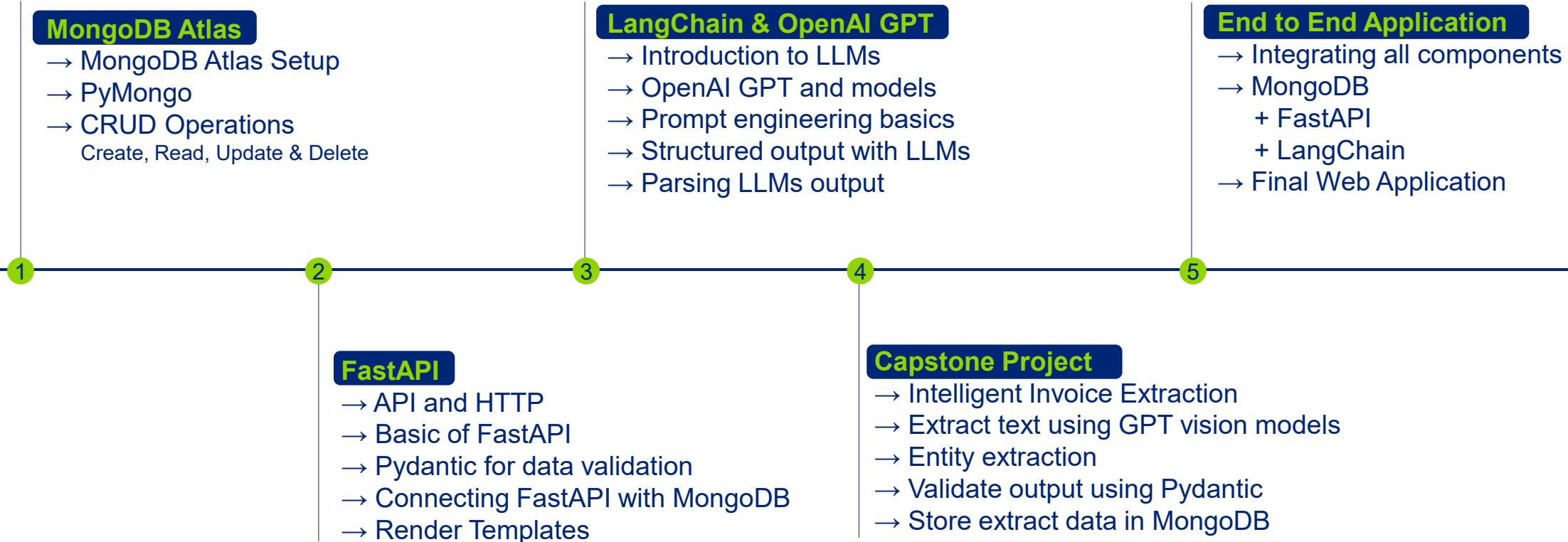
Architecture

Intelligent Document Extraction with OpenAI GPT & LangChain

→ Python, OpenAI API, Langchain, MongoDB, FastAPI



Road Map of the Course





Setting up Course

→ Pre-requisite: Python

Software

- Python
- IDE
 - Jupyter Notebook
 - VS Code

Setup

- Download Folder Structure
- Install Poetry → Dependency Management.
- Required Installations.
- Base Path & Create .env file → “secrets.env”

Download Folder Structure

→ GitHub: <https://github.com/vedicskill/udemy-intelligent-doc-openai>



Install Poetry

Python Packaging and Dependency Management



Install Libraries

Dependencies Libraries for the Course

Libraries

Modules	Packages
IDE, ENV	jupyter python-dotenv
Mongodb	pymongo[srv]
FasAPI	fastapi[standard] uvicorn httpx httpie requests
LangChain & OpenAI	langchain langchain-openai openai

Next | MongoDB Atlas

