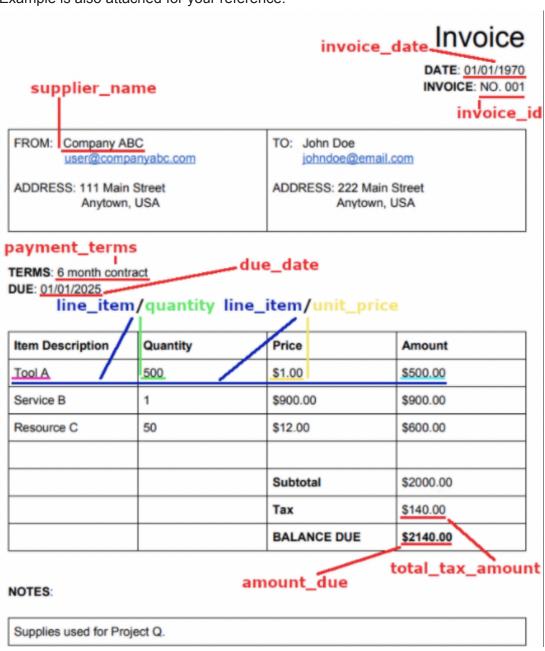Tasks:

Extract invoice number, invoice date, line items from invoice images. Also, the model structure should be scalable where we can train a model for any field and extract the key-value pair. Find sample data from the kaggle, huggingface, github. you can convert pdf files to image files.

Example is also attached for your reference.

invoice_date **Invoice**

DATE: 01/01/1970
INVOICE: NO. 001

invoice_id

supplier_name

FROM:	Company ABC	TO:	John Doe
	user@companyabc.com		johndoe@email.com
	ADDRESS: 111 Main Street		ADDRESS: 222 Main Street
	Anytown, USA		Anytown, USA

payment_terms

TERMS: 6 month contract
DUE: 01/01/2025 **due_date**

line_item/**quantity** **line_item**/**unit_price**

Item Description	Quantity	Price	Amount
Tool A	500	$1.00	$500.00
Service B	1	$900.00	$900.00
Resource C	50	$12.00	$600.00
		Subtotal	$2000.00
		Tax	$140.00
		BALANCE DUE	$2140.00

total_tax_amount

amount_due

NOTES:

Supplies used for Project Q.

Time Duration: 48 Hours.

If you have any queries, please feel free to contact us.