

Turning AI into API: Extracting Structured Data from Images with OCR and LLM

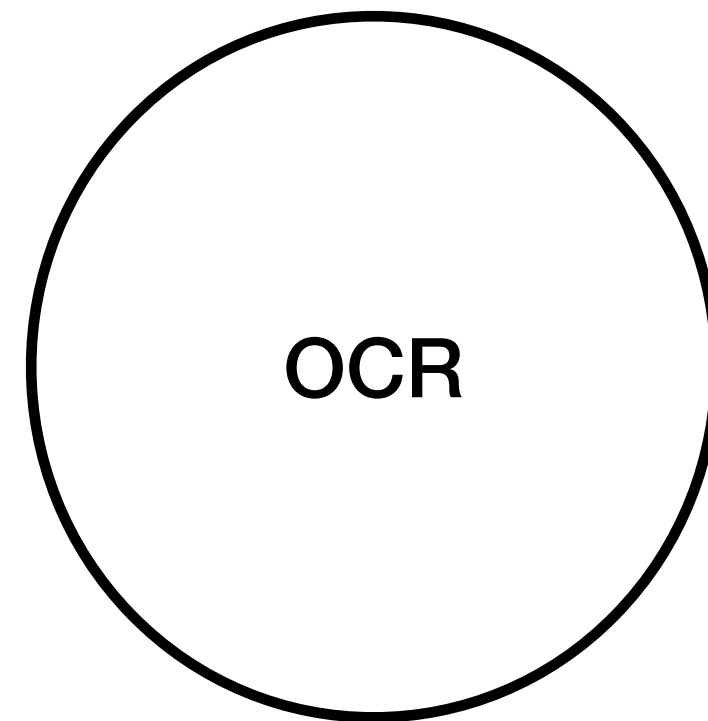
Vladimir Pesterev

What is OCR?

Optical Character Recognition

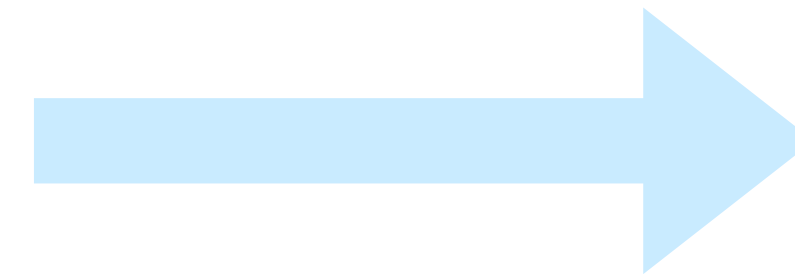
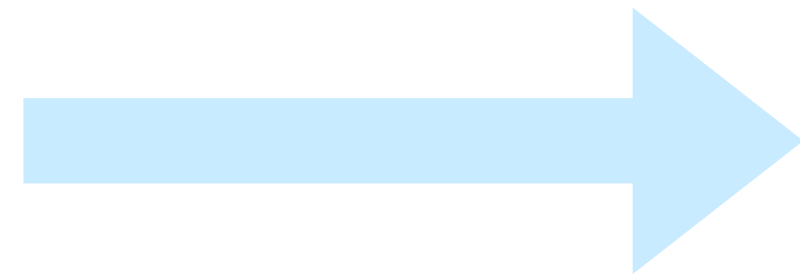
- Recognise text in images
- Amazon Textract or Google Cloud Vision API
- Tesseract.js

There is a problem

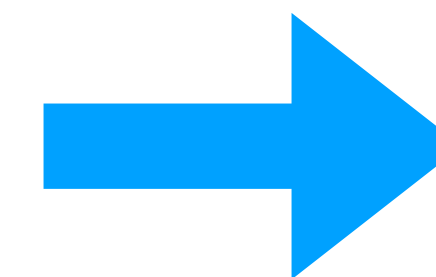
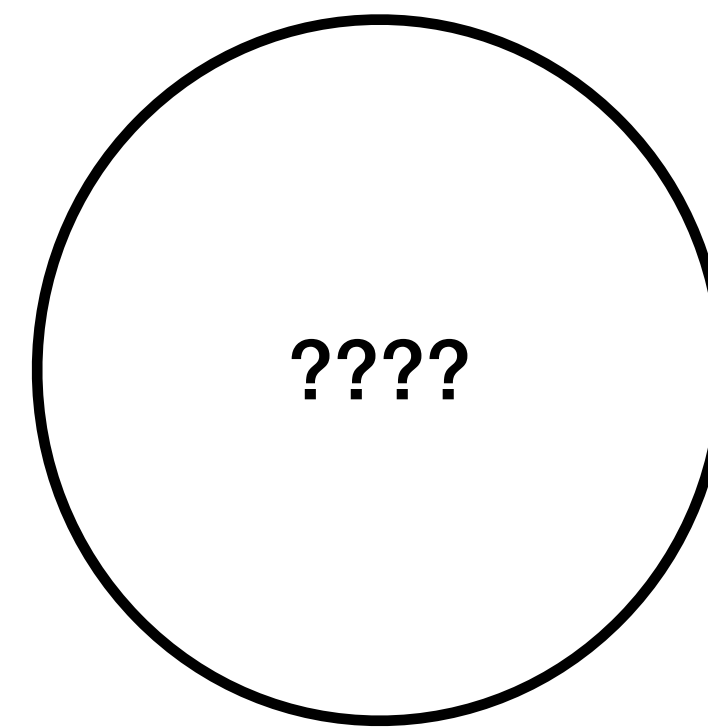
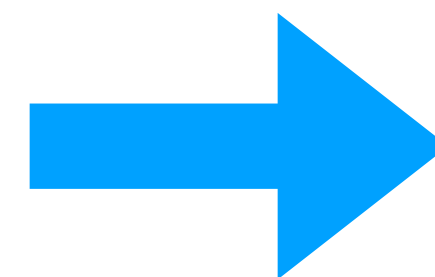
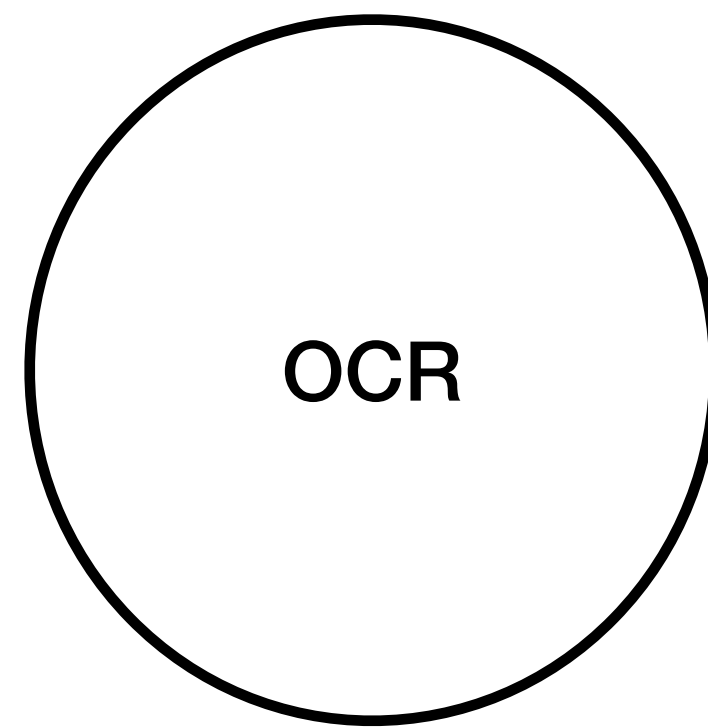
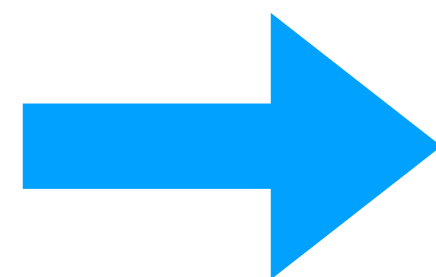


"RECEIPT ITEM 1 PRICE 1
ITEM 2 PRICE 2 TOTAL
\$100..."

There is a problem



```
"RECEIPT ITEM 1 PRICE 1  
ITEM 2 PRICE 2 TOTAL  
$100..."
```



```
{  
  "title": "RECEIPT",  
  "items": [  
    {  
      "title": "ITEM 1",  
      "price": "PRICE 1"  
    },  
    {  
      "title": "ITEM 2",  
      "price": "PRICE 2"  
    }  
  ],  
  "total": "$100"  
}
```

OpenAI API

- Provides interaction with ChatGPT models
- Function Calling is useful for many use cases
- JSON output mode
- Libraries: Python, JavaScript, .NET and more

Thank you

- <https://x.com/thechampysaur>
- conf42.smasher279@passinbox.com