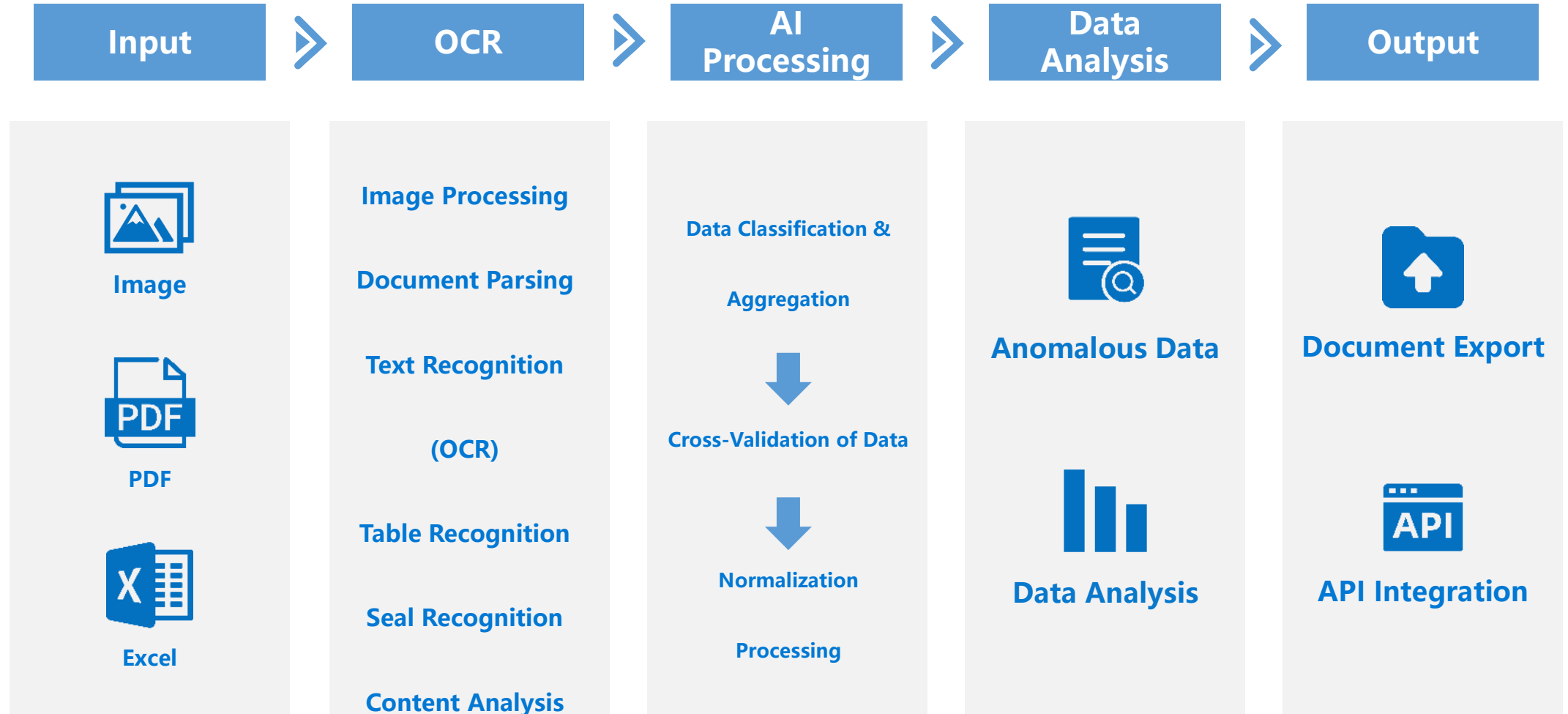


What is OceanX AI OCR?

- OceanX AI OCR is an AI-powered technology that automates the **capture**, **extraction**, **classification**, and **processing** of data from paper-based, digital, structured, semi-structured, and unstructured formats.
- By using technologies like **OCR**, **NLP**, and **ML**, Intelligent Document Processing reduces manual data entry, minimizes errors, and increases operational efficiency, allowing businesses to focus on higher-value tasks
- It leverages advanced AI models to handle complex document layouts, handwritten text, and multilingual content
- Designed for businesses that require fast, reliable, and secure document processing, OceanX AI OCR is deployable **on-premise** for full control over data

AI OCR Processing Workflow

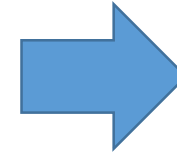


AI Capture

– Handwriting together with Printed Characters



致	:	Head of Department XYZ	
To	:	Head of Department XYZ	
名稱及標記 (物品性質及來源)	:	疑似白麵粉 suspected white flour –淨重量 ¹⁵⁰⁰ 1550 克放進 10 個塑膠袋裡面	
Description and Marking (Nature and source)	:	net weight ¹⁵⁰⁰ 750g in 10 plastic bags	
	:	如下所述印章編號 with seal no. as below	
	:	(Seal No.: 052404, 052405, 052406, 052407, 052408, 052409, 052410, 052411, 052412, 052413)	
包裝形式	:	Plastic bags with seal no. 膠袋上附有編號	
Packing type	:	Plastic bags with seal no. 膠袋上附有編號	
If Seal?	:	☒ YES ☐ NO	
總數量	:	包裝編號	052404, 052405, ⁰⁵²⁴⁰⁶ 052406,
Total Quantity	:	Packing Ref.	052407, 052408, 052409, 052410, 052411, 052412, ⁰⁵²⁴¹ 052413
送件人名稱	:	簽名	Wong Tai Hei
Delivered By	:	Signature	Wong Tai Hei
日期	:	時間	4 DEC 2025



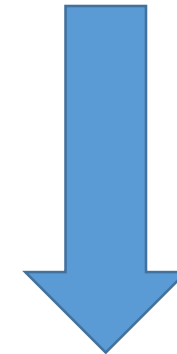
Prompt

选择文件 3 Sample For...g and Chi).pdf

Keywords to extract:

[致, 名稱及標記 (此欄位是多行文本), packing type, if seal, 總數量, 包裝編號]

upload



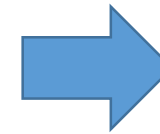
Result

```
{
  "content": {
    "致": "Head of Department XYZ",
    "名稱及標記 (此欄位是多行文本)": ["疑似白麵粉 suspected white flour -淨重量1500克放進 10 個塑膠袋裡面", "net weight 1500g in 10 plastic bags"],
    "packing type": "Plastic bags with seal no. 塑膠袋上附有編號",
    "if seal": "YES",
    "總數量": "10 nos.",
    "包裝編號": ["052404, 052405, 052406, 052407, 052408, 052409, 052410, 052411, 052412, 052413"]
  }
}
```

AI Capture – Capture of Signature Image



總數量	:	包裝編號	052404, 052405, 052406,
Total Quantity	10 nos.	Packing Ref.	052407, 052408, 052409,
			052410, 052411, 052412, 05
送件人名稱	:	簽名	
Delivered By	Wong Tai Hei	Signature	Wong Tai Hei
日期	4 DEC 2025	時間	
On		Time	4:30 p.m.
收件人名稱	:	簽名	
Received By	黃大喜	Signature	Wong Tai Hei
日期	-5 DEC 2025	時間	
On	-5 DEC 2025	Time	14:25
檔案編號	:		
Reference No.			
收取物品人士	:	簽名	
Returned To	龔竹荊	Signature	龔竹荊



Prompt

选择文件 3 Sample For...g and Chi).pdf

Keywords to extract:

[送件人簽名截圖, 收件人簽名截圖]

upload



Result

Signature Image result
in Base64 format

```
"content":{"送件人簽名截圖":  
"iVBORw0KGgoAAAANSUhEUgAAALMAAABICAIAAAAYl1N+AAAgAEIEQVR4Ae3BB5hdZb027ud537X23tP7TDK  
"收件人簽名截圖":  
"iVBORw0KGgoAAAANSUhEUgAAAL0AAAA4CAIAAAAnyG4MAAAGAEIEQVR4Ae3BB7yedX0+/uv6fO9nnj2yd8JI  
prompt_tokens:2091 completion_tokens:35
```

Wong Tai Hei Wong Tai Hei

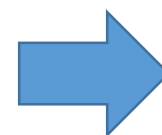
AI Capture – Understand the content to be captured



c) Dissolution
d) Uniformity of dosage units

3. Please contact Mr. CHAN Tai Man at 2123 2123.

[Handwritten signature]
21 MAY 2025



Prompt

选择文件 MEMO-1.Sa...with chops).pdf

Keywords to extract:

[聯絡人姓名, 聯絡人電話]

upload



Result

"content":{"聯絡人姓名": "Mr. CHAN Tai Man", "聯絡人電話": "2123 2123"}

prompt_tokens:2083 completion_tokens:34

AI Capture – Verification



第一部 申請人資料 (申請人必須是申請資助學生的父/母或是根據香港法例第13章《未成年人監護條例》下認可的監護人)

1. 中文姓名 KOON MBI MBI 2. 稱謂# ☒ A. 先生 ☐ B. 女士 ☐ C. 小姐

3. 英文姓名 KOON MBI MBI

4. 通訊地址 (請用英文填寫) Flat (室) 07 Floor (樓) 15 Block (座) B

大廈名稱 SUM YICK NATION

屋邨/村名稱 CHOI TUNG VILLAGE

街道名稱及號數 LUNG CHENG ROAD 112

分區 LUNG TAI SUI

地區 # ☐ 1. HK(香港) ☒ 2. KLN(九龍) ☐ 3. NT(新界) ☐ 4. OHK(香港境外)

5. 出生年份 1998

6. 香港身份證號碼 A1A 6010856

(如未能提供, 請在下方空格提供其他身份證明文件類別及號碼, 並附上副本)

其他身份證明文件類別: ☐ (請參閱《填寫及遞交綜合申請表格須知》第1.1段)

其他身份證明文件號碼:

7. 住宅電話號碼 23561234

8. 香港流動電話號碼 94538236 (學資處會以短訊形式發出各類通知, 必須填寫可以接收短訊的電話號碼)

9. 電郵地址

10. 申請人在1.4.2024-31.3.2025期間的婚姻狀況

☒ A. 已婚 (請在第二部提供配偶資料) ☐ B. * 離婚/分居/喪偶/未婚/其他 (請說明:)

(請提供有關證明文件副本, 而並留在第二部提供配偶資料)

11. ☐ 下一學年需要紙本申請表格 (註: 未有在方格內加上“✓”號, 將視為下一學年選用電子申請表格, 為便利申請及環保, 學資處鼓勵使用電子申請。)

12. 族裔 ☒ A. 華人 ☐ B. 巴基斯坦 ☐ C. 尼泊尔 ☐ D. 其他 (請註明:)

(註: 收集族裔的資料用作統計及研究用途, 並不會影響學生資助計劃綜合申請的處理及審核事宜。)

第二部 家庭成員資料和申請的資助計劃

A. 配偶

1. 中文姓名 沈俊威

2. 英文姓名 SHUM CHUN NAM

3. 出生年份 1998

4. 香港身份證號碼 MC 6010213

(如未能提供, 請在下方空格提供其他身份證明文件類別及號碼, 並附上副本)

其他身份證明文件類別: ☐ (請參閱《填寫及遞交綜合申請表格須知》第1.1段)

其他身份證明文件號碼:

5. 香港流動電話號碼 60800207

Indexing

申請人中文姓名 官美美

申請人稱謂 B. 女士

申請人英文姓名 KOON MBI MBI

通訊地址flat(室) 07

Floor(樓) 15

Block(座) B

大廈名稱 SUM YICK NATION

屋邨/村名稱 CHOI TUNG VILLAGE

街道名稱及號數 LUNG CHENG ROAD 112

分區 LUNG TAI SUI

地區 1. HK(香港)

申請人出生年份 1998

申請人香港身份證號碼 A1A 6010856

住宅電話號碼 23561234

香港物流電話號碼 94538236

申請人婚姻狀況 A. 已婚

族裔 A. 華人

配偶中文姓名 沈俊威

配偶英文姓名 SHUM CHUN NAM

配偶出生年份 1998

配偶香港身份證號碼 MC 6010213

香港流動電話號碼 60800207

AI Capture – Verification & Verifying Locator

全年總收入(\$) (包括花紅、津貼及兼職收入(不包括僱員強積金或公積金的強制性供款))	
薪金 (\$)	239400
營業盈利 (\$)	
薪金 (\$)	100000
營業盈利 (\$)	
薪金 (\$)	

Indexing	
申請人就業模式	全職
申請人職位	服務員
申請人薪金	239400
申請人配偶就業模式	兼職
申請人配偶職位	裁工比監
申請人配偶薪金	100000
總計	339400
賬戶持有人英文姓名	KOO N MEI MEI
申請人銀行戶口號碼	041-225-123456789
銀行名稱	創興銀行
申請人簽署	KOO N
申請人身份證文件號碼	MA6010455
申請人簽署日期	15.8.2025
申請人配偶簽署	EY)
申請人配偶身份證明文件號碼	MC6010213
申請人配偶簽署日期	15.8.2025

AI Capture – Verification & Verifying Locator



Indexing	
申請學生1中文姓名	沈学儿
申請學生1英文姓名	SHAN HOK YEE
申請學生1出生日期	2010年01月01日
申請學生1香港身份證	MC601019(1)
申請學生1 24-25期間的狀況	A. 在學
申請學生1 25/26學年就讀學校	Y.M.C.Y. Memorial Sch
申請學生1 25/26學年就讀年級	S3
申請學生1課程模式	A. 全日制
申請學生1申請資助計劃	需要
申請學生1中、小學或同等程度	3)書簿津貼 4)車船津貼
上網費津貼	不需要
申請人就業模式	
申請人職位	
申請人薪金	
申請人配偶就業模式	
申請人配偶職位	
申請人配偶薪金	

其他身份證明文件類別	其他身份證明文件號碼	2024-25期間的狀況	2025/26學年就讀學校或院校名稱	2025/26學年就讀年級
		# ☒ A. 在學	Y.M.C.Y.	S3

AI Capture – Verification & Verifying Locator

Indexing	
Block(座)	B
大廈名稱	SUN YICK MANSION
屋邨/村名稱	CHOI HUNG ESTATE
街道名稱及號數	LUNG CHENG ROAD 11
分區	WONG TAI SIN
地區	1.HK(香港)
申請人出生年份	1998
申請人香港身份證號碼	A1A 6010856
住宅電話號碼	23561234
香港物流電話號碼	94538236
申請人婚姻狀況	A.已婚
族裔	A.華人
配偶中文姓名	沈俊威
配偶英文姓名	SHUM CHUN NAM
配偶出生年份	1998
配偶香港身份證號碼	MC 6010213
香港流動電話號碼	60800207

AI Capture	
配偶	沈俊威
中文姓名	SHUM CHUN NAM
英文姓名	SHUM CHUN NAM
出生年份	1998
香港身份證號碼	MC 6010213
(如未能提供，請在下方)	
其他身份證明文件類別：	
其他身份證明文件號碼：	60800207
香港流動電話號碼@	
/A	

AI OCR vs Traditional OCR



OceanX AI OCR

- No form template needed.
- Deep learning models for context, layout, and language understanding
- OCR handwriting, and multilingual content effortlessly
- Continuously improves accuracy through machine learning
- Minimal manual intervention, ideal for diverse document formats

Traditional OCR

- Form templates matching
- Struggles with complex layouts, handwriting, and low-quality documents
- Need to setup a new template for new document type
- Requires manual adjustments for accuracy

Benefits of OceanX AI OCR



High Accuracy

- With AI enable, recognizes text, tables, and handwriting with far greater precision than traditional OCR.

On-Premise Solution

- Fully deployable on your infrastructure, ensuring data security, compliance, and control over sensitive information

Automatic Classification & Extraction

- Identifies document types and extracts key fields (names, dates, amounts) without manual work.

Supports Complex & Unstructured Documents

- Handles invoices, forms, contracts, scanned PDFs, and images—even with noise or distortion.

Improves Compliance & Data Quality

- Ensures consistent, governed data that integrates seamlessly with Back End applications & OceanX ECM.

Why Choose OceanX AI OCR?



- OceanX AI OCR is built on advanced AI models, ensuring cutting-edge accuracy and performance
- Its **on-premise deployment** ensures data privacy and security, making it ideal for industries with strict compliance requirements.
- **Document-Aware AI** — NO form templates needed, understands forms, tables & layouts.
- **Smart Data Extraction** — from text to structured data
- Designed for businesses that demand **accuracy, scalability, and flexibility**, OceanX AI OCR is trusted by organizations across finance, healthcare, logistics, and more.
- **Fast ROI** — Eliminates manual data entry, reducing processing time and operational expense.