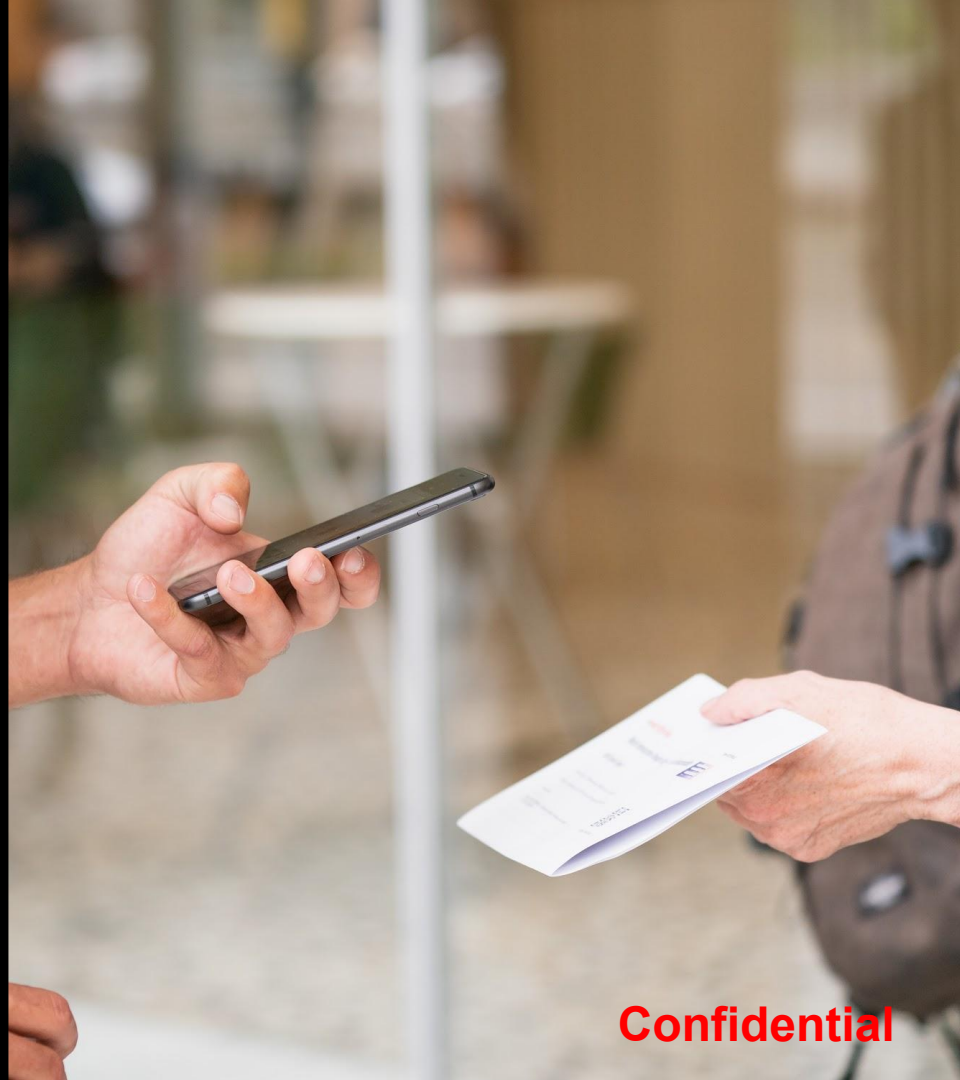


OURSKY

FormX.ai

Extract structured
information from
documents using AI



Confidential

FormX turns unorganised data returned from OCR engines to structured data in JSON, integrate with your software easily

From...

```
"人票易通知書 Cheque Deposit AVICE\nHSBC匯豐\n戶口資料\nAccount Information\nWP Monthly Parking Fee\nSTIC\n* * ANY PA克* * *\nD\n日期及時間\n銀行備考\nTVr201卷\n15:19:21\nBank Reference 3\n存入支票數目\nCheque\nscepted\n金額總數\n| Total Amount\nS3213\n客戶備註\nCustomer Reference\n20001:227555550282\n交易狀況\nTransaction Status\n接\n*1max, \n定心事29c\nm\n1:\n14: __\n14.iiiiiiiiiii:, \na.alist\n', '\nt\n_{\n:iti\nPial\n.\n:\n111", \ni\n"\ni=R\n二\n:bar;! \n中国在\n2030,32\n註\n:\n"。 \n.3.3.5\n1\n1\n"
```

Data got from an OCR engine

To...

```
{  
  "content": "\"WP Monthly Parking Fee",  
  "id": "fa6174b6-d530-4a63-887e-2d4122c18bb5",  
  "label": "account",  
  "languages": ["eng"]  
},  
{  
  "content": "$321.30",  
  "id": "e772d1fb-b8ed-4425-a10c-06d246cef7b0",  
  "label": "amount",  
  "languages": ["eng"]  
}
```

Structured data returned by FormX

Confidential

FormX unleashes the potential of the OCR technology

Problems of OCR

Traditional OCR engines scan and return blocks of text in images/PDFs. However, it does not know the meaning of the text and just return everything; For example, how do we know the date of transaction in a receipt?

Existing solutions

Currently developers will develop an algorithm to detect the block of text in the field by its position or semantic (e.g. if a text looks like a date or number). It is a manual process which is costly and error-prone.

Why FormX

FormX scans documents and returns structured data in JSON format. It also enables non-coders to define the position and the semantic of a field on a form via our Web Interface.

Confidential

FormX supports both fixed and dynamic format documents

Fixed Format

We use feature-based locational anchors to achieve higher accuracy.

Examples include:

- Passport or ID cards
- Business Registrations
- Driving Licenses and etc

Dynamic Format

We use layout models, key-value pairs, NER models to achieve a higher accuracy. Examples include:

- Mall receipts
- Invoices
- Bank statements and etc

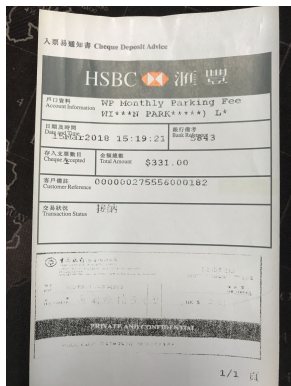
Optimisation

We provide customisation service if a higher accuracy is requested. Customers may provide around 100 samples for us to evaluate and we will revert our approach and induced service cost.

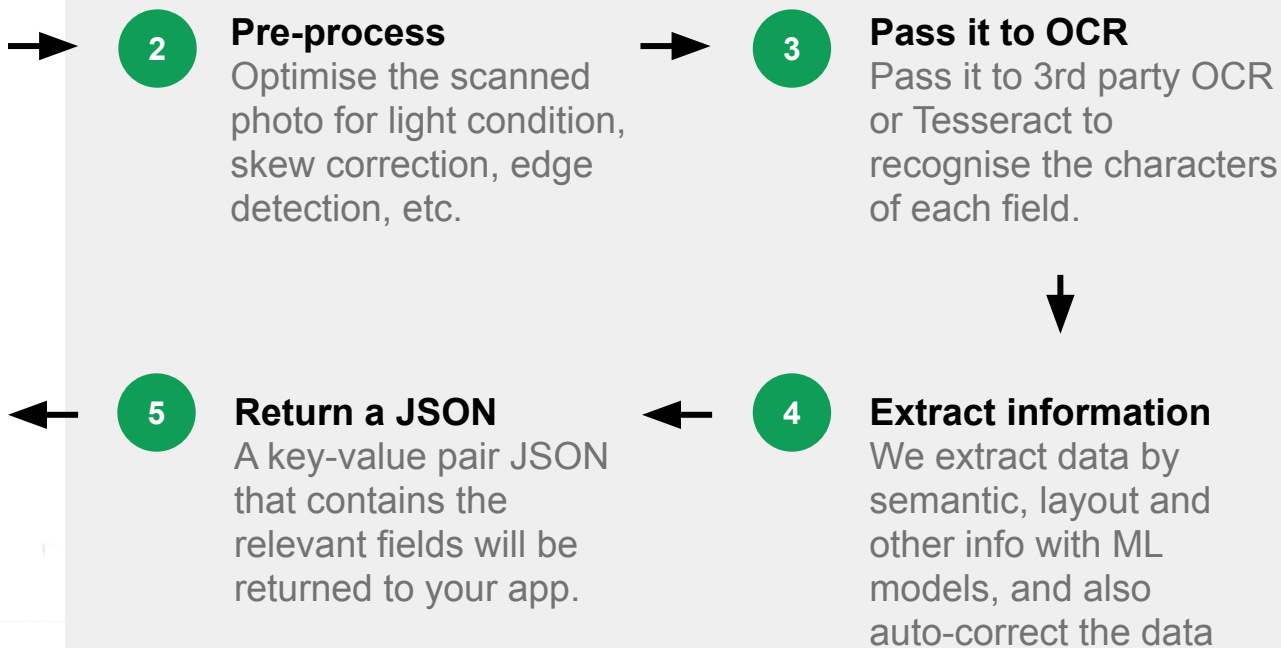
Confidential

How does it work?

- 1 Upload a photo**
Your users scan and upload a form via your app.



FormX



Key Features

1

Pre-built Document Parsers

While FormX works for all standardised forms, below are the forms we support by default:

- Hong Kong Business Registration
- Food Licenses / Travel Agency Licenses / Hawker Licenses
- Hong Kong ID Card
- Business cards
- Bank statements
- Motor vehicle license
- Consumer receipts

2

Create new Document Parsers

Processing documents beyond our pre-built parsers? No worry, you can define new parsers with our Web Portal.

3

Pre-Processing of Photos

FormX pre-process uploaded photos to improve the OCR accuracy. Photos are optimised for lighting condition, contrast, skew correction, edge detection, etc.

4

Auto Extraction to Specific Data Types

FormX includes native models for automatically extracting specific types of data from any formats of documents:

- Date and time
- Hong Kong Address
- Chinese Name
- Invoice Number
- Total amount (Usually for receipts)
- Product level information (Usually for receipts)
- Payment information (Usually for receipts)

Key Features (Cont'd)

5

Image recognition

Besides texts, FormX is able to recognise any images (Like a logo) by predefining those images in our portal as image tokens.

6

Classifier for Multiple Forms

If there are multiple formats from different forms, you can firstly define some unique key features (either by texts or images) to each type of forms, and all uploaded documents will be automatically assigned to the best-matched document parsers.

7

Multiple OCR engines support

FormX currently supports Google Vision API, Azure Computer Vision OCR and Tesseract.

8

Desktop Apps and Integrations

Besides JSON API, you can also use our desktop apps and G Sheet / Airtable integrations to batch process or create no-code solutions

Product Roadmap

We schedule new features to FormX.ai every quarter. Customers are welcome to provide insights or ideas of their feature wishlist.

By 2021Q3, FormX.ai will introduce new features/ enhancement including:

- 1 Fraud-checking to post-processed documents, e.g. Photoshopped
- 2 Enhancements for product level detail extraction of mall receipts

Confidential

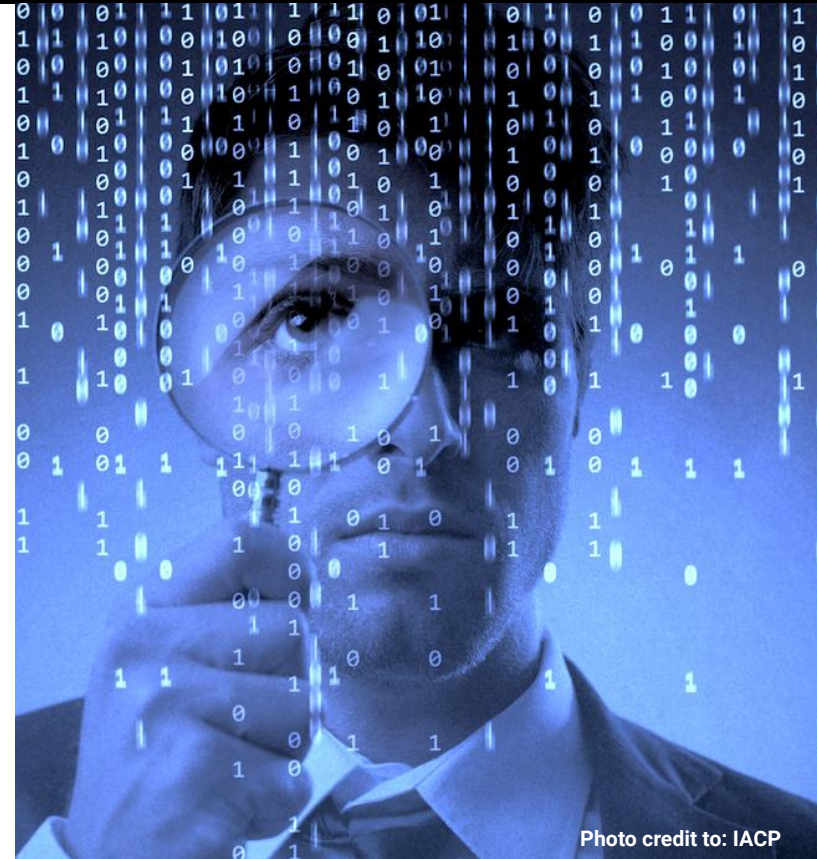
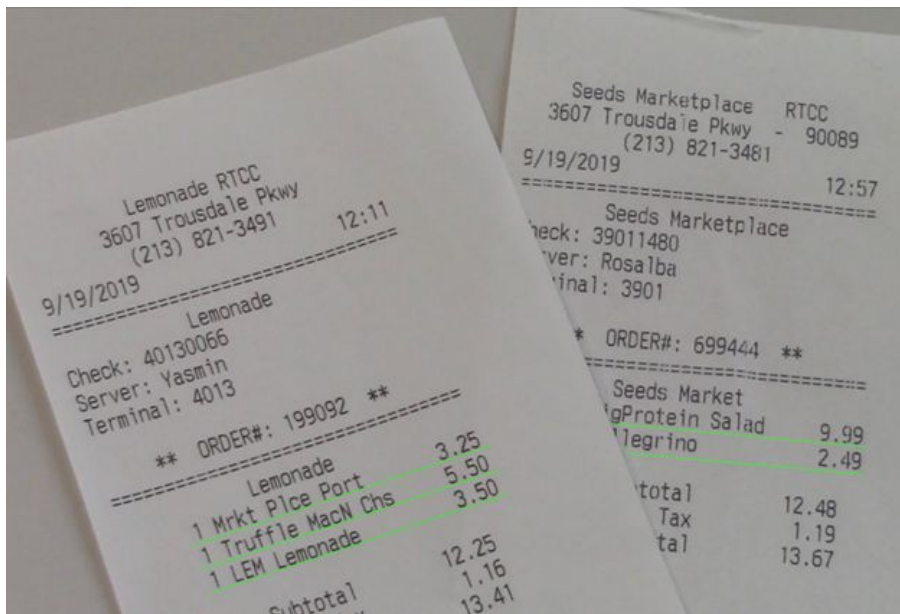


Photo credit to: IACP

Behind the scene - powered by Machine Learning



<https://blog.oursky.com/2020/04/06/formextractor-behind-the-scenes-how-artificial-intelligence-ai-is-used-to-extract-and-capture-data/>

Confidential

1

Line / Layout Estimation

Estimate and group words, identify relationships

2

Classifier and Auto-Correct Candidate

Built-in BoW Model to classifier and auto-correct amount, date, Chinese name, etc

3

Layout-specific ML Model

For dynamic format document, a specific model could be trained from around ~100 samples of document

Common Use Cases

How FormX.ai helps our Clients?

1

Enhance User Experience (UX)

Users no longer need to type input for applications or submission. All they need is take a photo on relevant documents, then FormX will extract the information and fill the form for the users.

2

Save Manpower

The colleagues who need to do manual input no longer need to type the data themselves. They can be the validator and check the FormX result only.

3

Speed Up Process

FormX shortens the time and minimize manual input error, Client can speed up lots of work process, application process to provide a better service for internal terms or customers.

Common use case

Improve customer engagement of your loyalty programme

Most property developers rely on loyalty programmes to retain their customers. Customers are rewarded loyalty points based on their spendings.

With FormX, customers can scan and upload their shopping receipts to your loyalty app and get the spending amount, date and other relevant information filled in for them. Making the data entry process convenient to your customers improve their engagement of your loyalty programme.





Common use case

Insurance claim - Reduce human error and cross-checking time

Money claims can be cost consuming since a huge amount of man-power is required for cross checking the receipt amount the customer inputs and the scanned receipt.

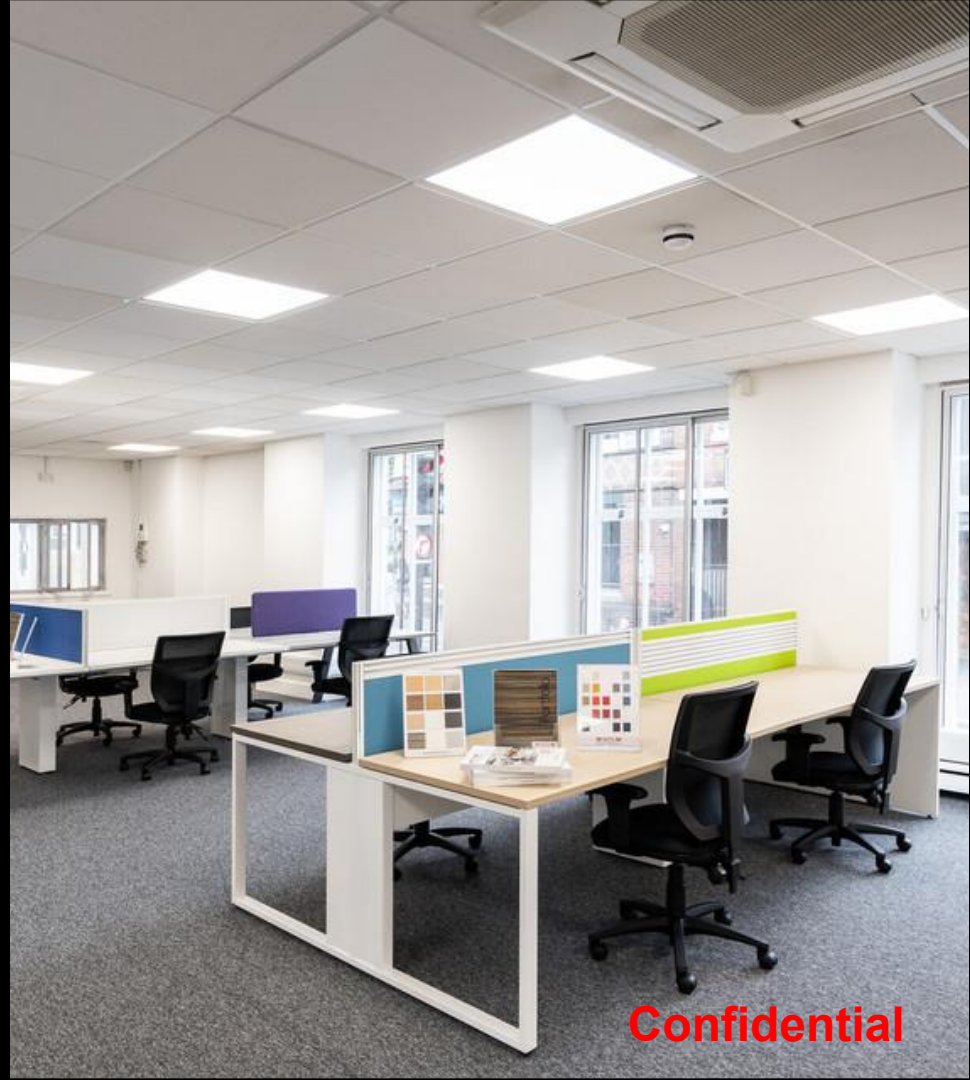
Instead of asking your customers to input the relevant information when they submit a claim, (which is also a tedious process for them), use FormX to extract the right information from the receipt to reduce human error and the cross-checking time.

Common use case

Office Automation - Collect key information from physical forms

Finding numbers from physical letters like invoices repeatedly can be very painful for an individual.

Try scanning Invoices from a physical scanner and FormX returns all key information from what you have previously defined.



Confidential

Case References

Accelerate Application Processes

Li Ka Shing Fund wanted to speed up the application process of the subsidy granted to small and medium businesses (SMBs) in Hong Kong.

How FormX helped:

- Extracted data such as license number and company name from
 - Business Registration (BR)
 - Food License
 - Hawker License
 - Travel Agency Licenses
- Deployed on production with accuracy improvement in 7 days
- BR recognition accuracy: 98%



Confidential

Accelerate HKSAR Government COVID-19 Relief Funds

Gov Covid-19 Retail Fund applied the same solution to enhance the user experience in form filling.

How FormX helped:

- Extracted data such as license number and company name from **Business Registration (BR)** documents

The image shows a screenshot of a web application for the HKSAR Government COVID-19 Relief Funds. The main heading is '同心抗疫' (同心抗疫). The form is titled '零售業資助計劃申請表格' (零售業資助計劃申請表格) and is at '步驟 1/4' (Step 1/4). The section '商業登記證' (Business Registration) is highlighted, with the instruction '請核實及確認自動填寫資料正確' (Please verify and confirm that the auto-filled information is correct). The form fields include:

- 您(申請者)的業務/法團名稱 (Your business/company name) with input boxes for 中文 (Chinese) and 英文 (English).
- 您的業務/分行名稱 (如有業務/分行名稱, 請填寫。) (Your business/branch name (if applicable, please fill in.)) with input boxes for 中文 (Chinese) and 英文 (English).
- 您的商業登記證號碼 (Your Business Registration Number) with a grid of input boxes.
- 您的商業登記證 生效日期 (Your Business Registration Effective Date) with dropdown menus for 年 (Year), 月 (Month), and 日 (Day).
- 您的商業登記證 屆滿日期 (Your Business Registration Expiry Date) with dropdown menus for 年 (Year), 月 (Month), and 日 (Day).

Confidential

4 relief funds by 3 different organizations uses FormX

Including Covid-19 Beauty Massage and Party room, D-Biz, and 2nd tranche of ESS

遙距營商計劃
Distance Business Programme

執行機構 Implementation Agent: 創新科技署 Innovation and Technology Commission
秘書處 Secretariat: hkpc 香港電力局

Part A – Particulars of the Applicant

Step 1/4

Upload of Documents
(Support file up to 5MB under format of jpg, jpeg)

Copy of a Valid Business Registration (“BR”) Certificate issued by the Inland Revenue Department which should be valid on the date of submission*

File Upload

Note 1: A valid Business Registration certificate issued under the Business Registration Ordinance (Cap. 310 of the Laws of Hong Kong) in respect of the business establishment named in the application and valid as of the application date is required.
Note 2: The image of the Business Registration certificate in JPEG format should be captured or scanned vertically and in its entirety, with details on the Business Registration certificate shown clearly. Failure to do so may render the application unsuccessful.

第二期「保就業」計劃
2nd tranche of Employment Support Scheme

立即申請 查看結果 計劃詳情 常見問題 守則及條款 已獲批名單

「保就業」計劃

申請資格
工資補貼額及計算方法
申請時間及資助發放安排
申請方法及審批所需資料
申請承諾
罰則
審查及監察機制
查詢

「保就業」計劃

香港特區政府在第二輪防疫抗疫基金下推出「保就業」計劃（計劃），透過向僱主的財政支援，協助他們支付員工的薪金，以保留可能曾被遣散的僱員。

計劃下向合資格的僱主提供的工資補貼分兩期發放：第一期補貼用以支付僱員在2020年9月至11月的工資；第二期補貼則用以支付僱員在2020年9月至11月的工資。「保就業」計劃大致完成審批及發放第一期補貼，並將於2020年8月31日至9月13日接受第二期申請。

任何參加計劃的僱主須承諾並保證 –
(一) 在接受政府工資補貼期間不會裁員；及
(二) 把政府工資補貼全數金額用於僱員工資。

此外，未有於第一期計劃領取一筆過7,500元資助的合資格自僱人士，亦可於2020年8月31日至9月13日期間提出申請。

1. 美容院、按摩院及派對房間資助計劃

網上看開申請進度
由即日起，申請者可按此連結輸入申請編號及商業登記號碼，查閱申請進度。

香港特別行政區政府（政府）在第二輪「防疫抗疫基金」下推出「美容院、按摩院及派對房間資助計劃」（本計劃），為受2019冠狀病毒疫情及政府各項防疫抗疫措施影響的合資格經營者提供援助。

1.1 美容院及按摩院申請資格

- 1.1.1 合資格的處所須為提供最少以下一項的服務作為主要業務：
 - (a) 皮膚及面部護理服務；
 - (b) 體重控制和纖體服務；
 - (c) 美甲服務；
 - (d) 髮身及/或髮型；
 - (e) 改善脫髮及/或護理；
 - (f) 按摩及/或水療護理；
 - (g) 足部按摩；及/或
 - (h) 對身體任何部位的其他美容、護理、美體及/或按摩服務。

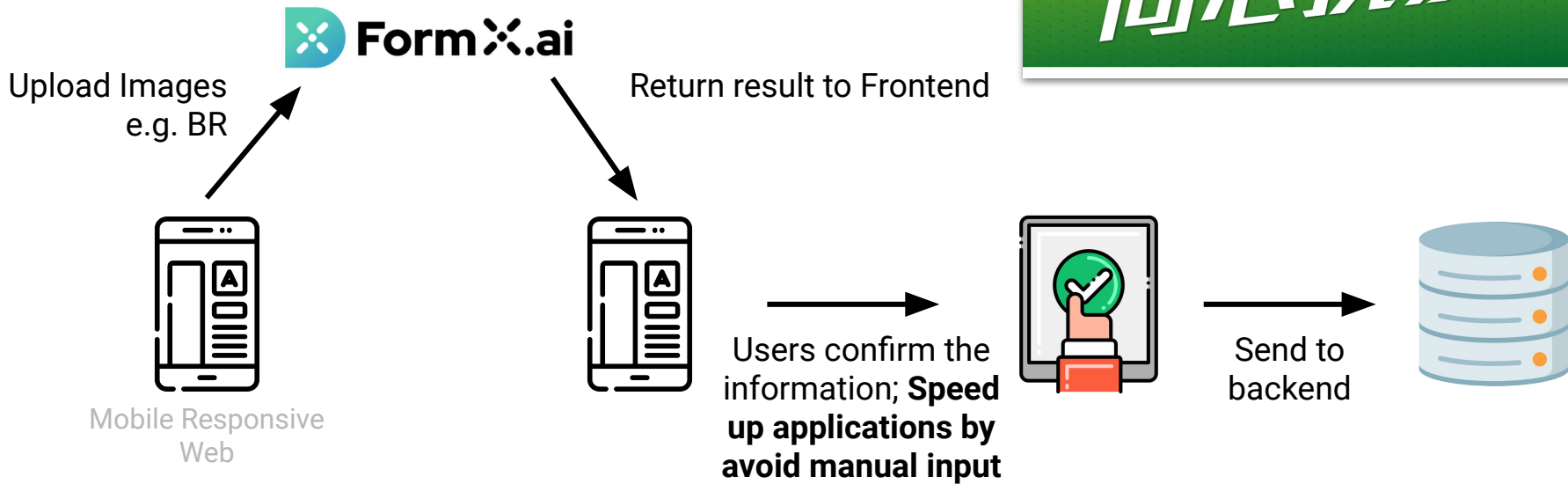
1.1 美容院及按摩院申請資格

1.1.1 合資格的處所須為提供最少以下一項的服務作為主要業務：

- (a) 皮膚及面部護理服務；
- (b) 體重控制和纖體服務；
- (c) 美甲服務；
- (d) 髮身及/或髮型；
- (e) 改善脫髮及/或護理；
- (f) 按摩及/或水療護理；
- (g) 足部按摩；及/或
- (h) 對身體任何部位的其他美容、護理、美體及/或按摩服務。

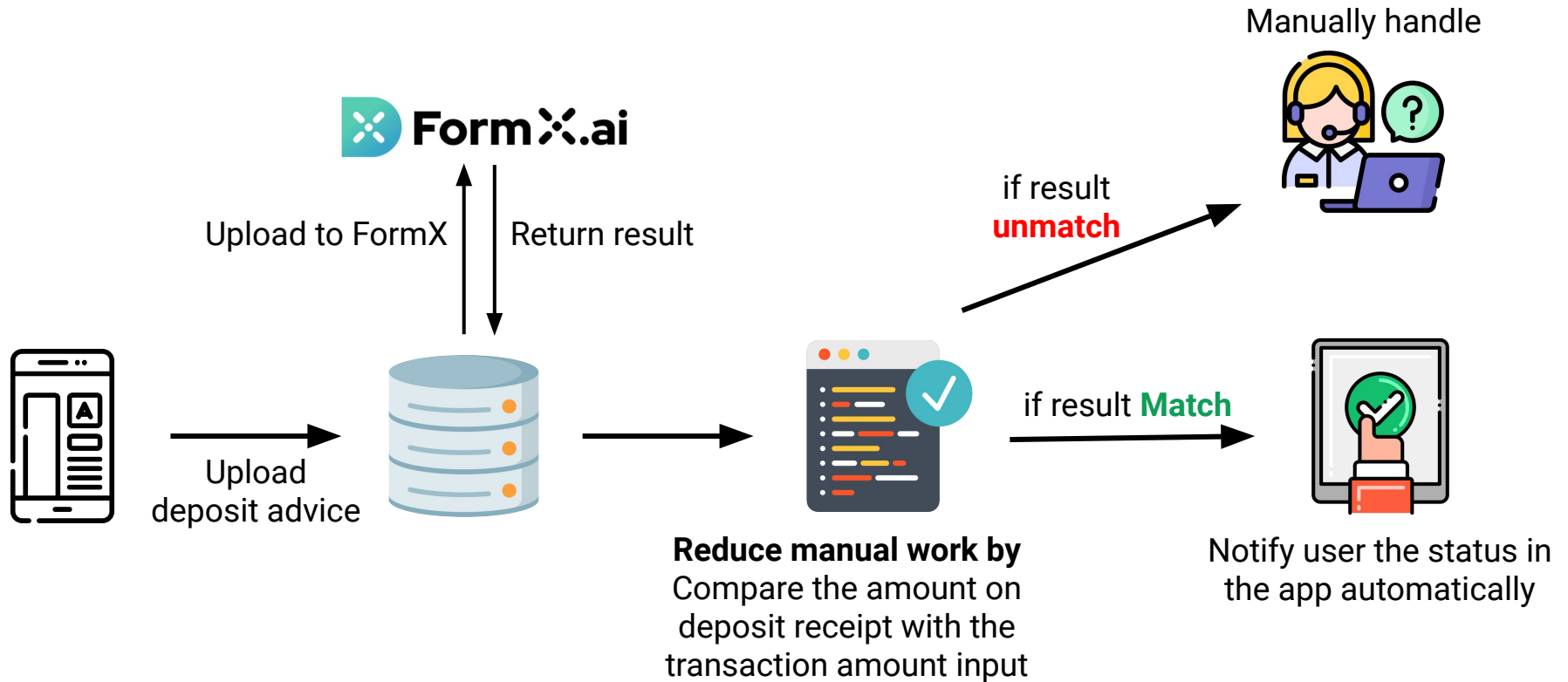
Confidential

How Li Ka Shing Foundation and HKSAR Gov enhance UX with FormX.ai?



Confidential

How Wilson Parking adopted FormX?



Shopping Mall Loyalty Program w/Receipts

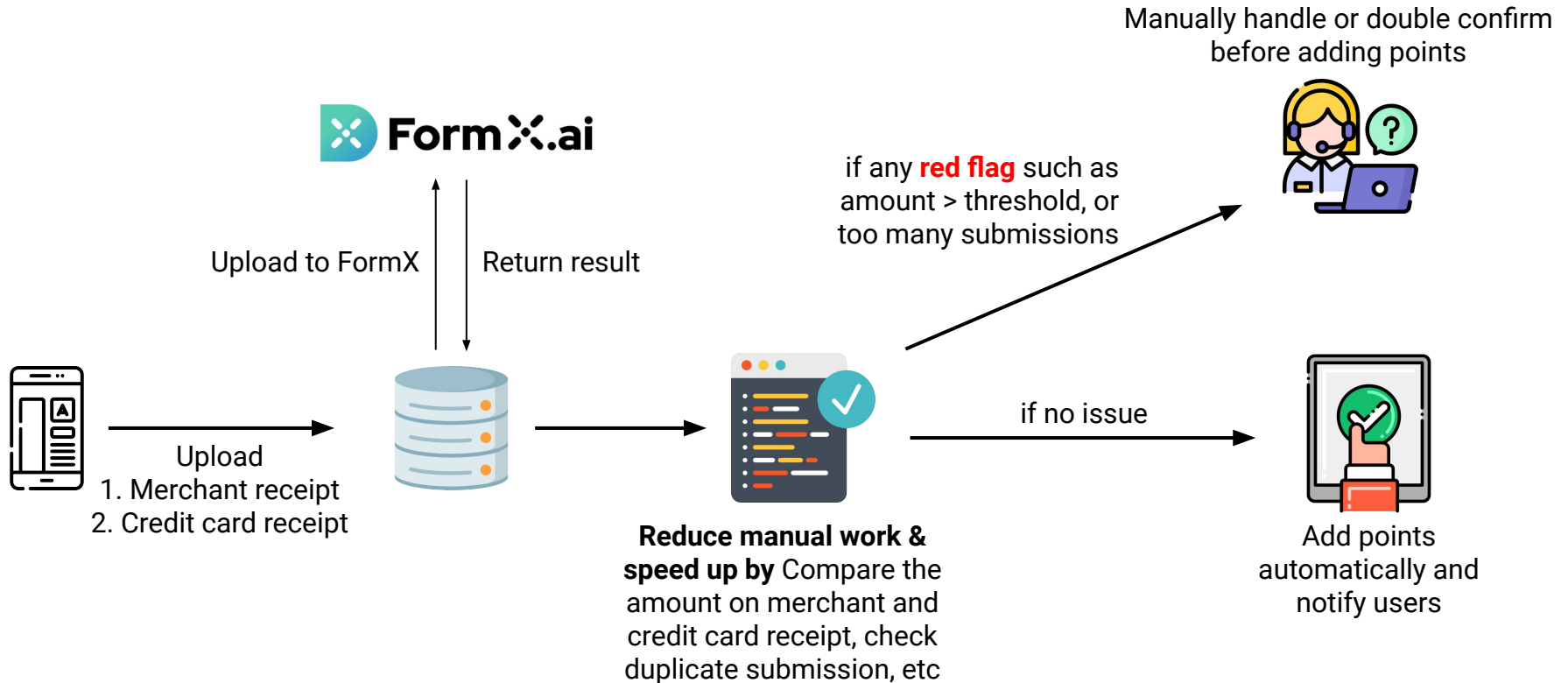
A property management company applied FormX to increase efficiency of data collection from mall receipts for loyalty program

How FormX helped:

- Extract data such as transaction date, time and total amount from mall receipts
- Extract credit card type
- Extract product line items
- Accuracy Reference
 - Date: 70% - 85%
 - Time: 75% - 90%
 - Amount: 68% - 87%



How Mall adopted FormX (receipt) for Loyalty Program?



Confidential

Automate data input from Motor Vehicle License

Link REIT, the largest Real Estate Investment Trust, applied this solution to increase efficiency of data collection from vehicles.

How FormX helped:

- Extracted data such as vehicle registration number and expiry date from Motor Vehicle License (MVL)



API Integration

HTTP Endpoint

POST `https://worker.formextractorai.com/extract`

FormX will isolate extraction fields from the uploaded image then perform OCR on these fields.

Form ID can be submitted as multipart/form-data or through a header. Image can be uploaded as is or through an image_url.

For uploading image directly, it can be uploaded in the request body or via multipart/form-data.

On the other hand, if you want to specify an image_url, it can be submitted via a header or multipart/form-data.

cURL Example

```
curl -X POST \  
https://worker.formextractorai.com/extract \  
-H 'Content-Type: image/jpeg' \  
-H 'X-WORKER-FORM-ID: REPLACE-YOUR-FORM-ID-HERE' \  
-H 'X-WORKER-TOKEN: REPLACE-YOUR-WORKER-TOKEN-HERE' \  
--data-binary "@/path/to/query/image.jpg"
```

We designed a single endpoint API for easier integrations with other applications.

Return JSON format

```
{  
  "status": "ok",  
  "form_id": "7bc280ba-277e-4d51-a187-7a7e7e9dafb1",  
  "fields": [  
    {  
      "name": "dateOfCommencement",  
      "region_id": "acacf0de-c259-4fa9-873b-902b2ac7e281",  
      "type": "script",  
      "value": "15/01/2018"  
    }  
  ],  
}
```

Background of the Product Owner - OURSKY

Leader in Tech

Founded by 3 x computer engineers since 2009

Over half of headcount are full time developers

12+ year Experience in agile development, Elite player

Refresh in-house tech stack every quarter

Leader in app UX

In-house software UI/UX team with a design lead with 12 years of relevant experiences

Handled different APP modernisation projects with Enterprise using a startup perspective

Recognised

Over 50 people company

Best of 2013 & 2017 from Apple App Store; ICT Award 2014 - Award of the Year

Worldwide Customers: U.S., U.K., Germany, Italy, Singapore, Australia, Japan and more.

Confidential

Who BELIEVES in OURSKY?

ArtStack



NEXT DIGITAL



HKPC

Snaptee

TATA COMMUNICATIONS



A.S. Watson Group
A member of CK Hutchison Holdings



Week in China



Request a demo at
info@oursky.com

Confidential