

Incident date	2020-05-27
Incident time	14:00
LE27 date	2020-05-27 (same day)
LE29 date	
Building type	Shopping centre (Private)
Victim involved	4
Victim admitted to hospital	1
Death count	0

#6 ✓ 3

#14 ✓ 2

#1 ✓ 2

00:00 / 00:02

Response #15 Accepted

Previous Job

Next Job



Aric Fung

ANNOTATIONS

ANNOTATION LABEL

- head
- ✓ wings

STEPS

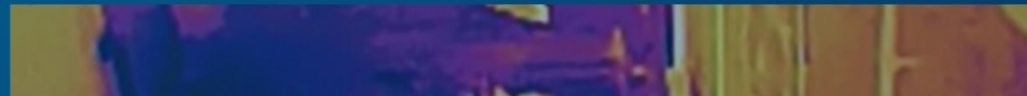
- ✓ Click to Select a Corner
- Click to Select the Opposite Corner

Adjust the Box

cancel



All you need to start your AI journey.



love crystals

淨。晶

攝位查詢

或可... 請認明... 詳情請... 詳情請... 詳情請...

Enhance Workplace Efficiency with
Automatic speech recognition (ASR), Natural language Processing (NLP)
and Optical Character Recognition (OCR)

Company Profile

Datax AI Solutions

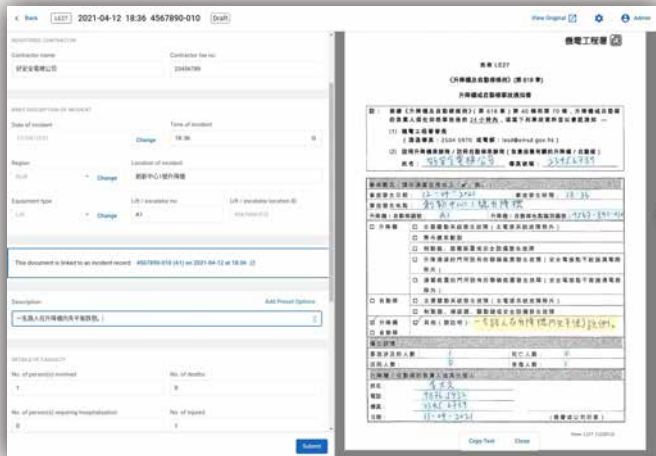
- ❖ A local tech-heavy company with 20 team members
- ❖ Specialized in delivering OCR, Speech-to-text, NLP & computer vision solutions
- ❖ Gained >30 gov & corporate clients in 4 years since establishing the AI solutions unit

Some of our Clients with Commercially Launched Solutions

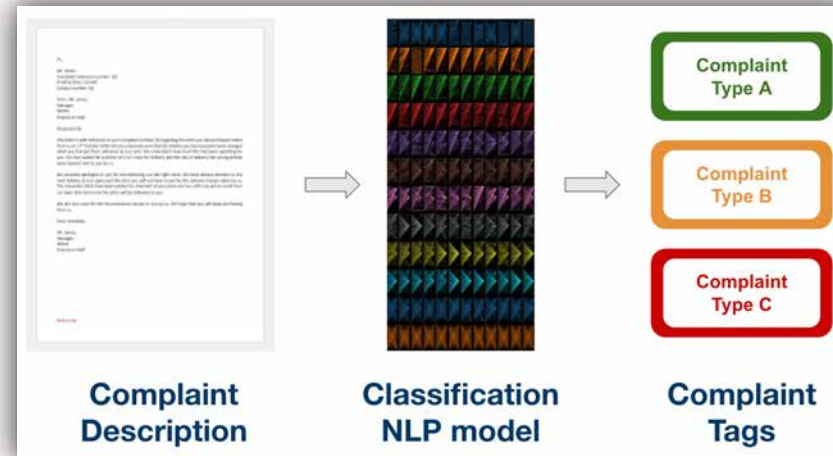


Datax AI Solutions

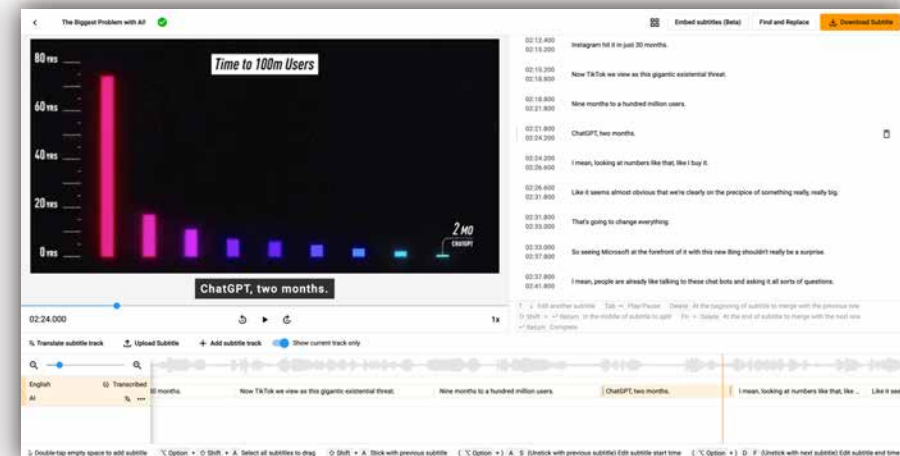
Automation by NLP & OCR



Handwriting OCR & Document Extraction



GPT and LLM based-NLP solutions



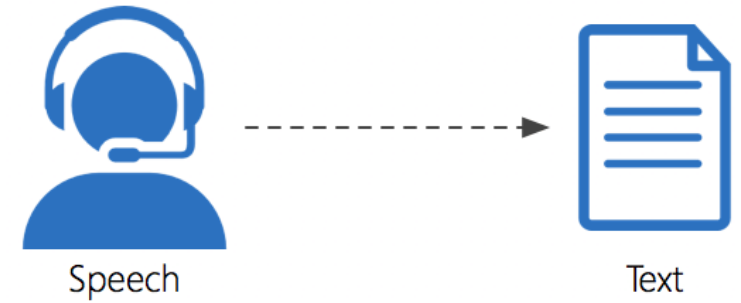
Multiple Language Speech-to-text (ASR)

Language Solutions

- ❖ Speech-to-text (ASR), subtitling
- ❖ Translation
- ❖ Classification
- ❖ Sentiment analysis
- ❖ Summarization
- ❖ Entity extraction
- ❖ GPT-related applications

Datax Speech-to-text (ASR)

Audio transcription & video subtitling



1. Accurate & fast

- ❖ Superior transcription accuracy compared to similar software, with average below 5%-word error rate (95% accuracy)
- ❖ Process time short to 1/6 of audio length
- ❖ Capable to handle audio with significant background noise

2. Versatile

- ❖ Supports 80+ common languages, including direct generation of Written Chinese or Verbal Chinese from Cantonese audio
- ❖ Customizable to fit business's daily workflows
- ❖ Seamless integration with downstream NLP solutions for multiplying business value generated
(Enables subtitling, complaint classification, voice information tagging, call monitoring, meeting summarization etc.)

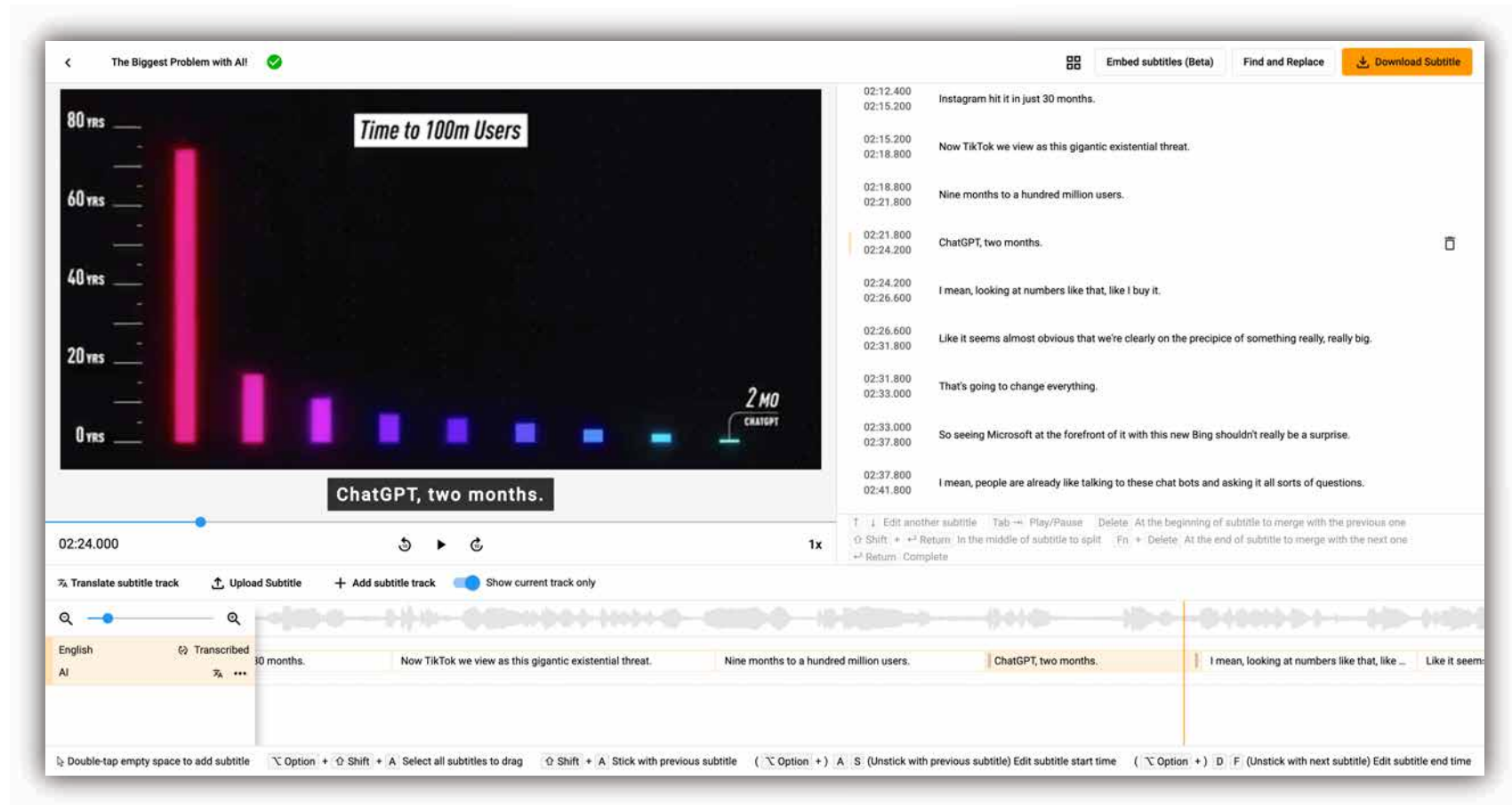
3. Secure & reliable

- ❖ Support offline, on-premise implementation and usage
- ❖ Online version build on Google, Azure and Amazon cloud services with file encryption

Subanana – Leading Cantonese Subtitling Platform

Audio transcription & video subtitling

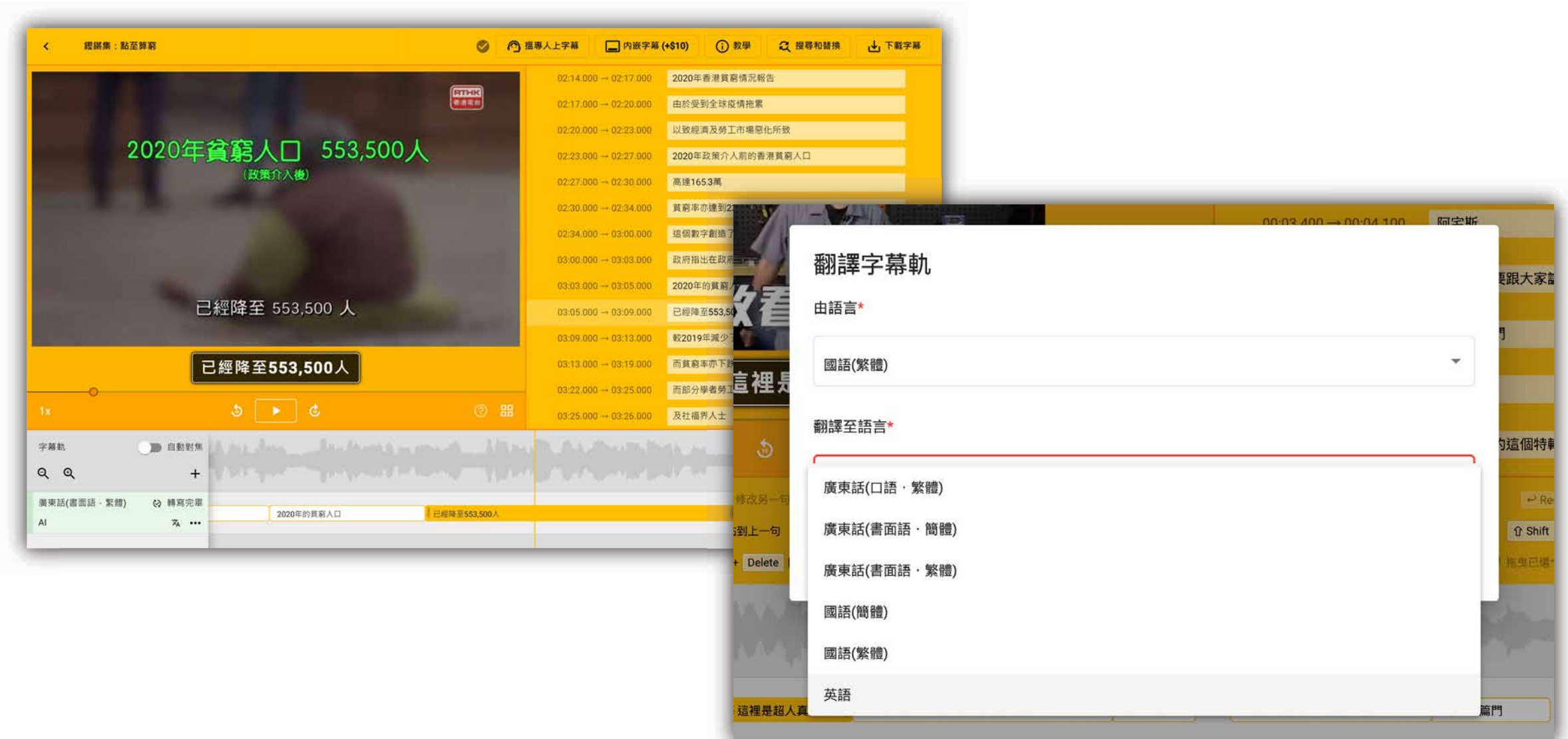
- ❖ Launched for public access since 2021
- ❖ Trusted by 200+ YouTube channels, media groups & production companies
- ❖ Provide subtitle for avg. 1000+ hours of video per month



<https://subanana.com/plus/>

Subanana – Leading Cantonese Subtitling Platform

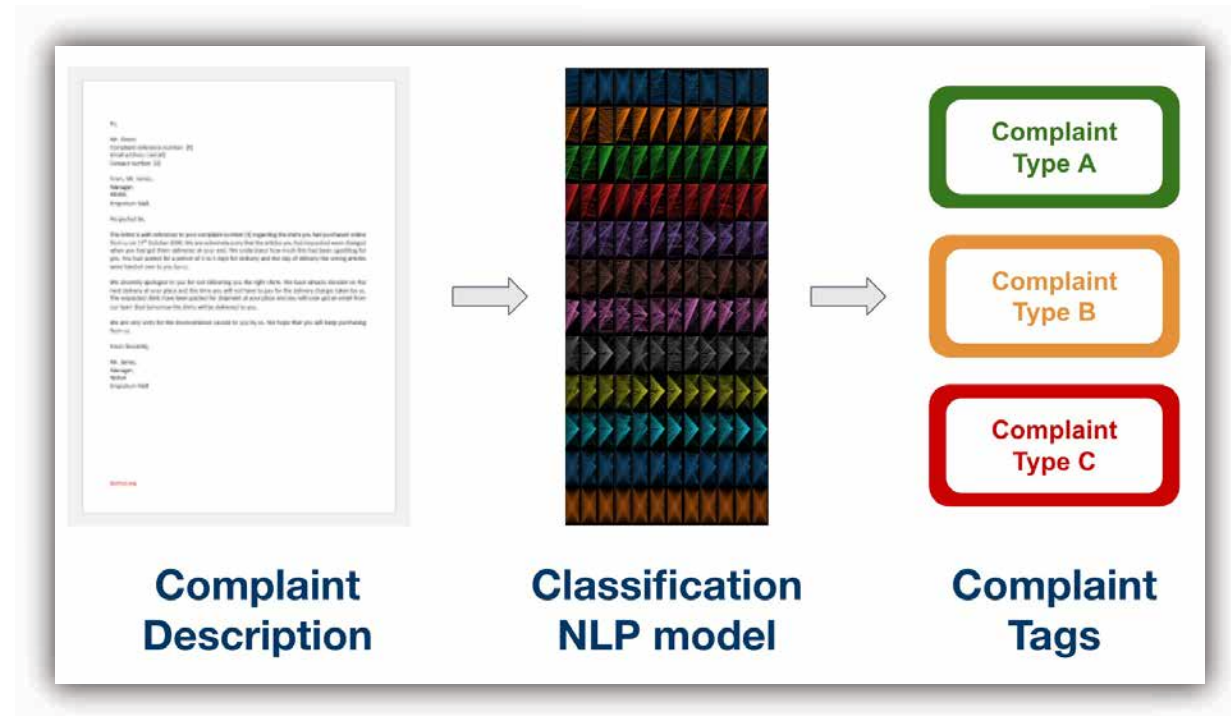
Audio transcription & video subtitling



Complaint Classification & Reporting

Project Showcase - Serving commercial bank & public transportation company

- ❖ NLP AI models to read & understand customer enquiries, not only based on keywords
 - ❖ Language supported: Traditional & Simplified Chinese, English, & a mix of them
- ❖ The NLP system can assist staffs to:
 1. Identify the core problem by adding complaint tags
 2. Distribute complaint to different division in a bank
 3. Identify customer sentiment & urgency
 4. Suggest appropriate replies or solutions



Complaint Classification & Reporting

Project Showcase - Serving commercial bank & public transportation company

Complaints [Bulk Tag Adjustments] [Import Tag Adjustments] [Export] [Delete]

Filters & Columns

Created At: 2022-02-20 → 2022-02-28 Imported At: 2022-02-20 → 2022-02-28

Application For: select Product: select Type: select Sub Type: select

Theme Tags: tag-1 x Add Process Tags: tag-1 x Add Role Tags: tag-1 x Add

Keyword (Whitelist): whitelist x type to search by keyword whitelist

Keyword (Blacklist): blacklist x type to search by keyword blacklist

Columns:

Available: 1/14 items

- CEMS ID
- Created Date
- Channel
- Product
- Type
- Sub Type
- Status

Selected: 9 items

- CEMS ID
- Created Date
- Channel
- Product
- Type
- Sub Type
- Status

[Apply]

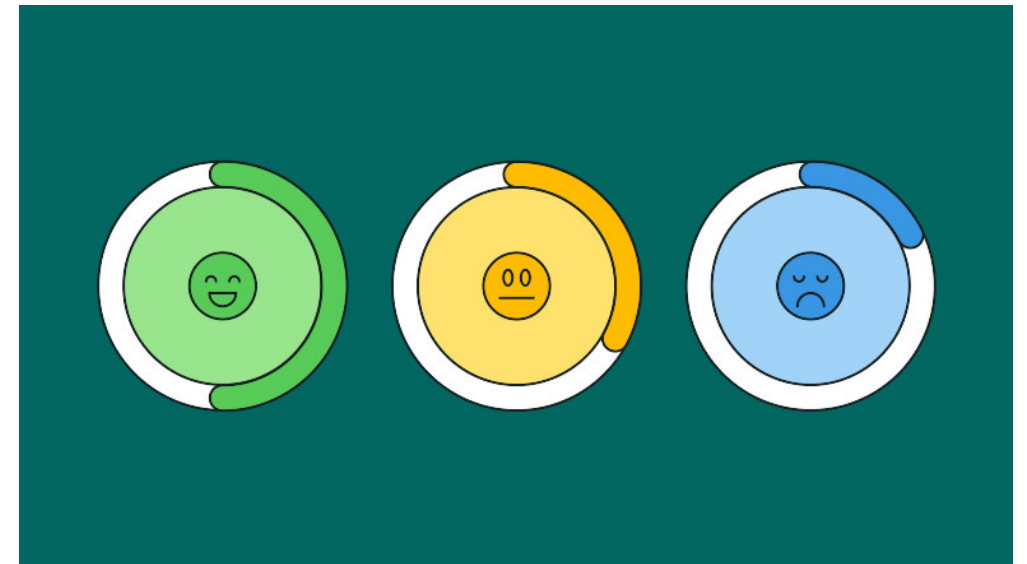
<input type="checkbox"/>	CEMS ID	Created Date	Channel	Application	Product	Type	Sub Type	Theme	Process	Role	View
<input type="checkbox"/>	BC_202204261618	2022-04-26	website	Investment	Fixed...	Product...	Credit...	late-finance-charge	aml-cdd	cdm	

load more

Sentiment Analysis

Project Showcase - Serving commercial bank & public transportation company

- ❖ NLP AI models to identify sentiment and **give reason to its prediction**
 - ❖ Positive, neutral, negative, mixed or scale-based analysis
 - ❖ Ability to breakdown the customer inquiry & provide sentiment analysis if 2 or more topics are mentioned



Project Showcase - Housing patrol recording analysis

Speech-to-text with Key Entity Extraction

Model: medium

Language: Chinese

URL (YouTube, etc.):

Upload Audio: 0:00 / 0:09

Microphone Input: Record from microphone

Initial Prompt (override Format):

VAD: silero-vad-skip-gaps

Download

cp-垃圾vtt3174im4a-sub.srt	98.0 B	Download
cp-垃圾vtt3174im4a-sub.vtt	98.0 B	Download
cp-垃圾vtt3174im4a-transcript.tx...	64.0 B	Download
cp-垃圾vtt3174im4a-result.json	791.0 B	Download

Transcription: 現在走到一場15號位 有些垃圾在這裏 找人來收拾

Entities: [LOC]: 一場15號位, [ITEM]:, [ISSUE]: 有些垃圾在

Segments: WEBVTT, 00:00.122 --> 00:08.122, 現在走到一場15號位 有些垃圾在這裏 找人來收拾

Flag

Download

f-雜物foa2o903m4a-sub.srt	61.0 B	Download
f-雜物foa2o903m4a-sub.vtt	61.0 B	Download
f-雜物foa2o903m4a-transcript.txt	27.0 B	Download
f-雜物foa2o903m4a-result.json	547.0 B	Download

Transcription: 21樓南翼大堂 有雜物

Entities: [LOC]: 21樓南翼大堂, [ITEM]:, [ISSUE]: 有雜物

Segments: WEBVTT, 00:00.000 --> 00:03.000, 21樓南翼大堂 有雜物

Flag

VAD: silero-vad-skip-gaps

Clear Submit

Document & Text Solutions

- ❖ AI-OCR for hand-writings & typed text
- ❖ Structural document extraction (Forms)
- ❖ Semi-structured document extraction (Invoices, bank statements)
- ❖ Data redaction
- ❖ Image text recognition

Next-Gen OCR for Commercial Use

Turn hand-written physical documents into digitized & structured data

1. Extract info from scanned docs with OCR technology

- ❖ Handwriting & printed texts
- ❖ Traditional & Simplified Chinese, English, Numbers
- ❖ Structured & semi-structured documents
- ❖ Best for Hong Kong corporates

2. Include document pre-processing

3. Provide flexible output approaches

- ❖ Direct output as CSV file
- ❖ API integration through JSON, XML file
- ❖ System & database integration

The image shows a side-by-side comparison of a scanned document and its digitized, structured representation. On the left, a web interface displays the extracted data in a form. On the right, the original scanned document is shown with the extracted text overlaid, demonstrating the accuracy of the OCR process.

REGISTERED CONTRACTOR

Contractor name: 好安全電梯公司 | Contractor fax no.: 23456789

BRIEF DESCRIPTION OF INCIDENT

Date of incident: 12/04/2021 | Time of incident: 18:36

Region: KLV | Location of incident: 創新中心1號升降機

Equipment type: Lift | Lift / escalator no.: A1 | Lift / escalator location ID: 4567890-010

This document is linked to an incident record: 4567890-010 (A1) on 2021-04-12 at 18:36

Description

一名路人在升降機內失平衡跌倒。

DETAILS OF CASUALTY

No. of person(s) involved: 1 | No. of deaths: 0

No. of person(s) requiring hospitalization: 0 | No. of injured: 1

機電工程署

表格 LE27
(升降機及自動梯條例) (第 618 章)
升降機或自動梯事故通知書

註：根據《升降機及自動梯條例》(第 618 章)第 40 條和第 70 條，升降機或自動梯的負責人須在知悉事故後的 24 小時內，填寫下列事故報告並以書面通知 —

(1) 機電工程署署長
(透過傳真：2504 5970 或電郵：lesd@emsd.gov.hk)

(2) 註冊升降機承辦商 / 註冊自動梯承辦商 (負責保養有關的升降機 / 自動梯)

姓名：好安全電梯公司 | 傳真號碼：23456789

事故簡述 (請在適當空格加上「✓」號)：

事故發生日期：12-04-2021 | 事故發生時間：18:36

事故發生地點：創新中心1號升降機

升降機 / 自動梯編號：A1 | 升降機 / 自動梯地點識別編號：4567-890-010

升降機

- 主要驅動系統發生故障 (主電源系統故障除外)
- 吊鉤或索鉤鬆脫
- 制動器、緩速裝置或安全設備發生故障
- 升降機通道的門所設有的聯鎖裝置發生故障 (安全電掣點不能接過電路除外)
- 滿載裝置的門所設有的聯鎖裝置發生故障 (安全電掣點不能接過電路除外)

自動梯

- 主要驅動系統發生故障 (主電源系統故障除外)
- 制動器、梯級鎖、驅動鏈或安全設備發生故障

升降機

- 其他 (請註明)：一名路人在升降機內失平衡跌倒。

自動梯

獨立詳情

事故涉及的人數：1 | 死亡人數：0

送院人數：0 | 受傷人數：1

升降機 / 自動梯的負責人或其代理人

姓名：李大明

電話：98765432

傳真：23456789

日期：12-04-2021 | (好安全公司印章)

Sample OCR result validation & input UI

Next-Gen OCR for Commercial Use

Digitization for business documents

❖ Structured data extraction from scanned or captured documents, e.g.

1. Business Registration (BR)
2. Invoices
3. Bank statements (various banks supported)
4. ID cards, passports
5. Receipts
6. And more...

	Invoice #1 Chat Horn Engineering Ltd 23 04 HKD 89.830.00
	Invoice #2 Chat Horn Engineering Ltd Date HKD 19.743.00
	Invoice #3 Chat Horn Engineering Ltd HKD 64.70700
	Invoice #4 Chat Horn Engineering Ltd
	Invoice #5 Chat Horn Engineering Ltd Date HKD 6.00

Company	Chat Horn Engineering Ltd
Beneficiary	Chat Horn Engineering Ltd
Invoice Number	117509..
Invoice Date	23 04
Account Number	CH80106296
Amount	HKD 89.830.00

Next-Gen OCR for Commercial Use

Competitive advantages of Datax OCR

1. Self-developed OCR models

- ❖ AI trained with >10 million hand-writing data in Chinese & English, **best for Hong Kong corporates.**
- ❖ Specifically fine-tuned for recognizing & extracting common data fields, incl.
 - ❖ Addresses
 - ❖ Date & time
 - ❖ Numbers
- ❖ Can be deployed **on premise**, private cloud or public cloud

2. Highly accurate after AI training (Verified by official model testing in gov projects)

- ❖ **>85% for hand-written texts**
- ❖ **>95% for typed texts**

Project Showcase 1 – EMSD LERIS System

Contractor name: 好安全電梯公司
 Contractor fax no.: 23456789

BRIEF DESCRIPTION OF INCIDENT

Date of incident: 12/04/2021
 Time of incident: 18:36

Region: KLW
 Location of incident: 創新中心1號升降機

Equipment type: Lift
 Lift / escalator no.: A1
 Lift / escalator location ID: 4567890-010

This document is linked to an incident record: [4567890-010 \(A1\) on 2021-04-12 at 18:36](#)

Description: 一名路人在升降機內失平衡跌倒。

DETAILS OF CASUALTY

No. of person(s) involved: 1
 No. of deaths: 0

No. of person(s) requiring hospitalization: 0
 No. of injured: 0

表格 LE27
 《升降機及自動梯條例》(第 618 章)
 升降機或自動梯事故通知書

註：根據《升降機及自動梯條例》(第 618 章)第 40 條和第 70 條，升降機或自動梯的負責人須在知悉事故後的 24 小時內，填寫下列事故資料並以書面通知 —

(1) 機電工程署署長
 (透過傳真：2504 5970 或電郵：lesd@emsd.gov.hk)

(2) 註冊升降機承辦商 / 註冊自動梯承辦商 (負責保養有關的升降機 / 自動梯)
 姓名：好安全電梯公司 傳真號碼：23456789

事故概況 (請在適當空格加上「✓」號。)

事故發生日期：12-04-2021 事故發生時間：18:36
 事故發生地點：創新中心1號升降機
 升降機 / 自動梯編號：A1 升降機 / 自動梯地點識別編號：4567-890-010

<input type="checkbox"/> 升降機	<input type="checkbox"/> 主要驅動系統發生故障 (主電源系統故障除外)
	<input type="checkbox"/> 懸吊纜索斷裂
	<input type="checkbox"/> 制動器、超載裝置或安全設備發生故障
	<input type="checkbox"/> 升降通道的門所設有的聯鎖裝置發生故障 (安全電接點不能接通電路除外)
	<input type="checkbox"/> 運載裝置的門所設有的聯鎖裝置發生故障 (安全電接點不能接通電路除外)
<input type="checkbox"/> 自動梯	<input type="checkbox"/> 主要驅動系統發生故障 (主電源系統故障除外)
	<input type="checkbox"/> 制動器、梯級鏈、驅動鏈或安全設備發生故障
<input checked="" type="checkbox"/> 升降機	<input checked="" type="checkbox"/> 其他 (請註明)：一名路人在升降機內失平衡跌倒。
<input type="checkbox"/> 自動梯	

傷亡詳情

事故涉及的人數：1 死亡人數：0
 送院人數：0 受傷人數：1

升降機 / 自動梯的負責人或其代理人

姓名：李大文
 電話：98765432
 傳真：23456789
 日期：12-04-2021 (簽署或公司印章)

Project Showcase 2 – LandsD Performance Appraisal Report OCR Tool

JAFPARS Appraisal Spellcheck (1).pdf Export as .docx Clear System settings Logout

Copy this page 1 / 6 Annotation

Suggested Texts All

Filter Any Copy All

PERFORMANCE APPRAISAL

for the period 1.7.2020 to 31.12.2020

Personal data provided in this performance appraisal report are to facilitate the appraisal. Access to the completed report is limited to the appraisee

(2) Full name of officer (surname first):
Mr Chan Tai-man (陳大文)

(3) Month and year of birth:

Page 1 Raw Corrected

Suggestion: Full Name of Officer

Mr Chan Tai-man

Page 1 Raw Corrected

Suggestion: Present Post

LI I/DLO/TW&KT 1,7.2020

Page 3 Raw Corrected

Suggestion: Name of AO

Wong Tai Man

Project Showcase 3 – LandsD Job Application Form (G.F.340) OCR System

JAFPARS | JAF Restricted (Staff) admin

Recruitment exercises > test on 2023-02-20 > Application 000013

Application #000013 Mark as reviewed

Reset view Jump to form field Go

香港特別行政區政府職位申請表 Application Form for Employment with the Government of the Hong Kong Special Administrative Region	
申請職位名稱 Title of job applied for	地政總署 Department
A部 Section A 個人資料 Personal Particulars	
英文姓名 Name in English	SIU MEI
中文姓名 Name in Chinese	黃小美
出生日期 Date of birth	1 1 1990
性別 Sex	<input checked="" type="checkbox"/> Female

Search detected text All Corrected Raw

< 1 / 6 >

[Save changes](#)

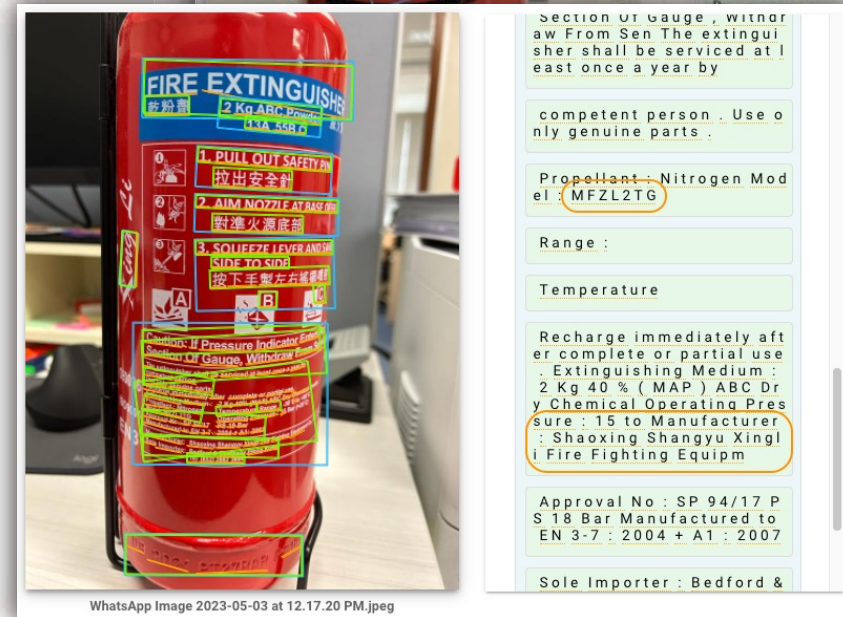
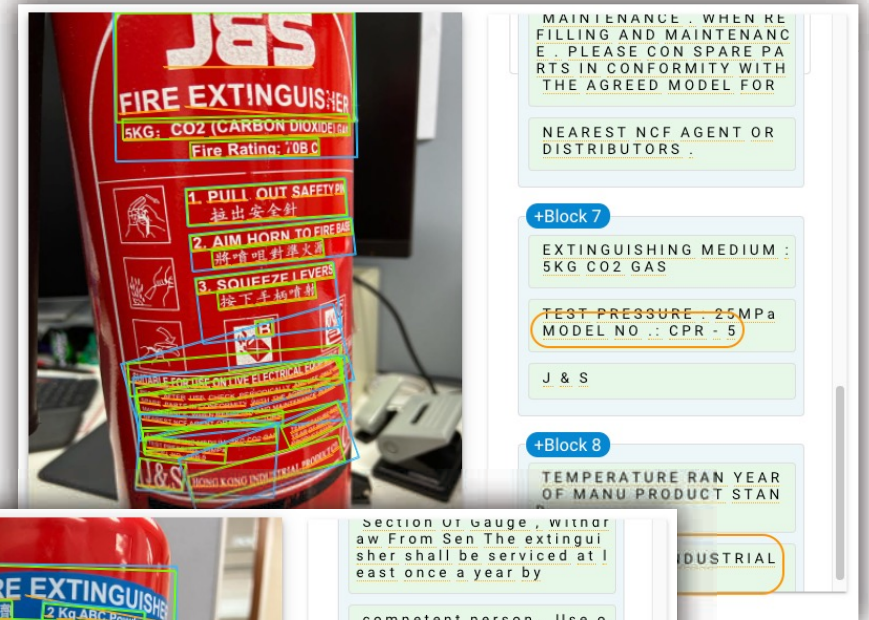
Project Showcase 5 – FSD Port-E App

Image Text Recognition

❖ Implement an app module for public to search for approved portable fire equipment

1. Allow users to capture their own portable fire equipment by camera
2. Use OCR solutions to extract key info from the captured image
3. Compare and classify the captured portable fire equipment with the approved database of FSD
4. Display and rank most similar approved portable fire equipment for the public checking
5. Demo video:

https://www.youtube.com/shorts/Dx6hbr_A108



WhatsApp Image 2023-05-03 at 12.17.20 PM.jpeg

Project Showcase 6 – Car License Recognition

- ❖ Locate and label the position of car license
- ❖ OCR included engine to extract the text and turn into digital format
- ❖ Integrate with corresponding data libraries to enhance extraction accuracy
- ❖ Demo Video:
<https://www.youtube.com/watch?v=vjeG3nAN8co>





datax.io

