Kevin Chen
陳宇成

ADDRESS

Rm. 2, 8F., No.12, Ln. 222, Jinlong Road, Neihu district, Taipei, Taiwan

CONTACT

+886 986 308 886

Kevinjc.tw@gmail.com

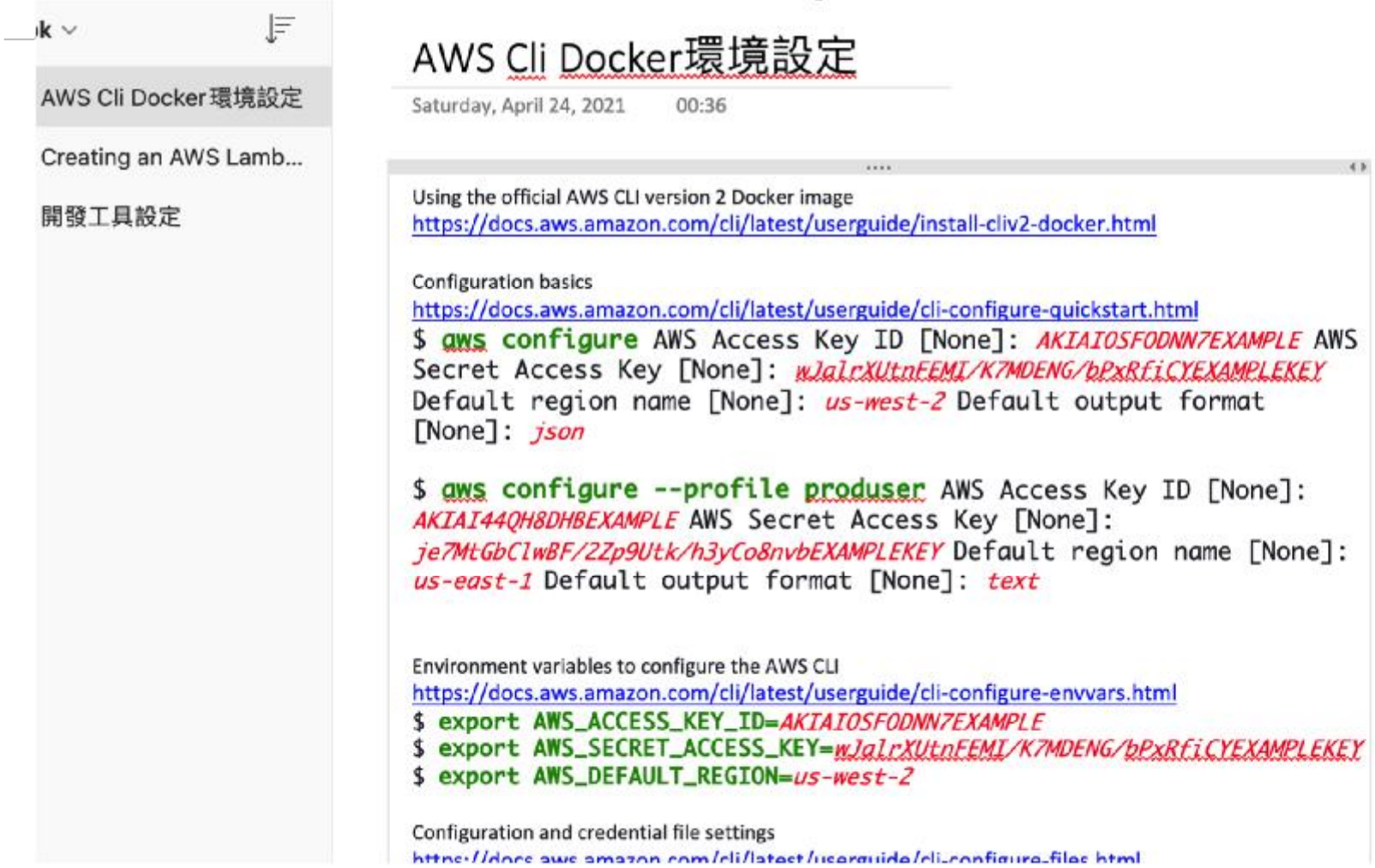
Uchiha_kevin

Kevinjc_tw

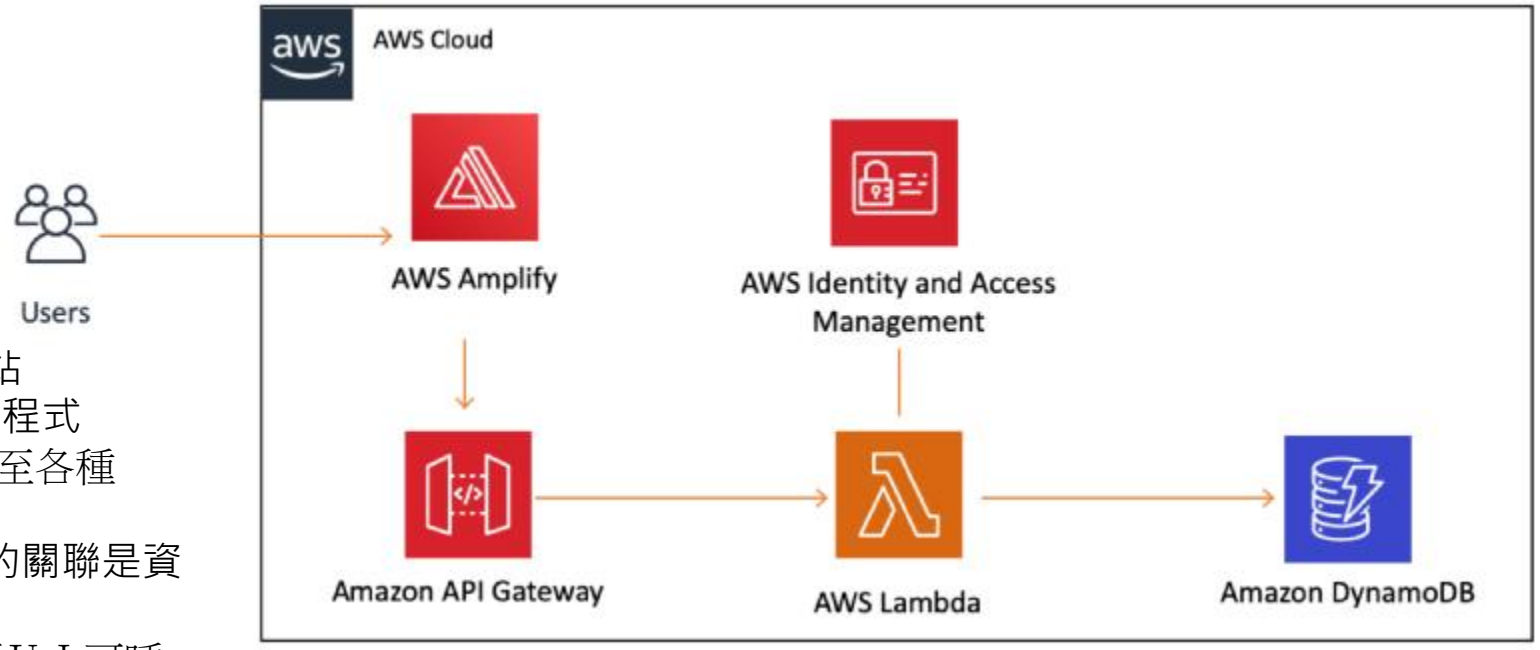
陳宇成簡介

- 具有25年以上豐富的資訊技術經驗與經營導向的管理專業
- 前瞻的實務經驗在資訊科技反映出穩固質量的理解能力、治理部門、專案管理、系統整合與架構設計
- 具製造業的AI數據分析、機器學習管理與分析能力
- 採用VS.net帶領與實務開發MS BI應用之各類專案類型：
  - 資料整合以SSIS開發ETL資料匯流至 ODS或DW、以SSRS開發企業ERP所需報表、 以SSAS開發資料多維度模型
  - 採用Power BI 桌面版設計資料來源由一般資料庫、SSAS與SSRS2的管理儀表板
- 微軟Azure雲端技術整合、Azure ML、IoT Hub/Edge應用，應用即時資料流/資料庫設計與規劃(Kafka+ Nifi, MS-SQL/Oracle)

近日ＡＷＳ開發實作使用PyCharm



近日ＡＷＳ建立基本 Web 應用程式



- AWS Amplify: 靜態網頁網站
- AWS Lambda: 無伺服器API程式
- AWS API GW: API統一連結至各種 Lambda
- Amazon DynamoDB: AWS 的關聯是資料庫服務
- AWS 互動式Web app: 網頁ＵＩ可呼叫API GW連結的Lambda funciton

近日ＡＷＳ開發實作使用PyCharm

ook

AWS Cli Docker環境設定

Creating an AWS Lamb...

開發工具設定

開發工具設定

Saturday, April 24, 2021 00:50

Key tasks for the AWS Toolkit for JetBrains
https://docs.aws.amazon.com/toolkit-for-jetbrains/latest/userguide/key-tasks.html#key-tasks-install

Installing the AWS SAM CLI on macOS
https://docs.aws.amazon.com/serverless-application-model/latest/developerguide/serverless-sam-cli-install-mac.html

/bin/bash -c "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/master/install.sh)"
brew --version

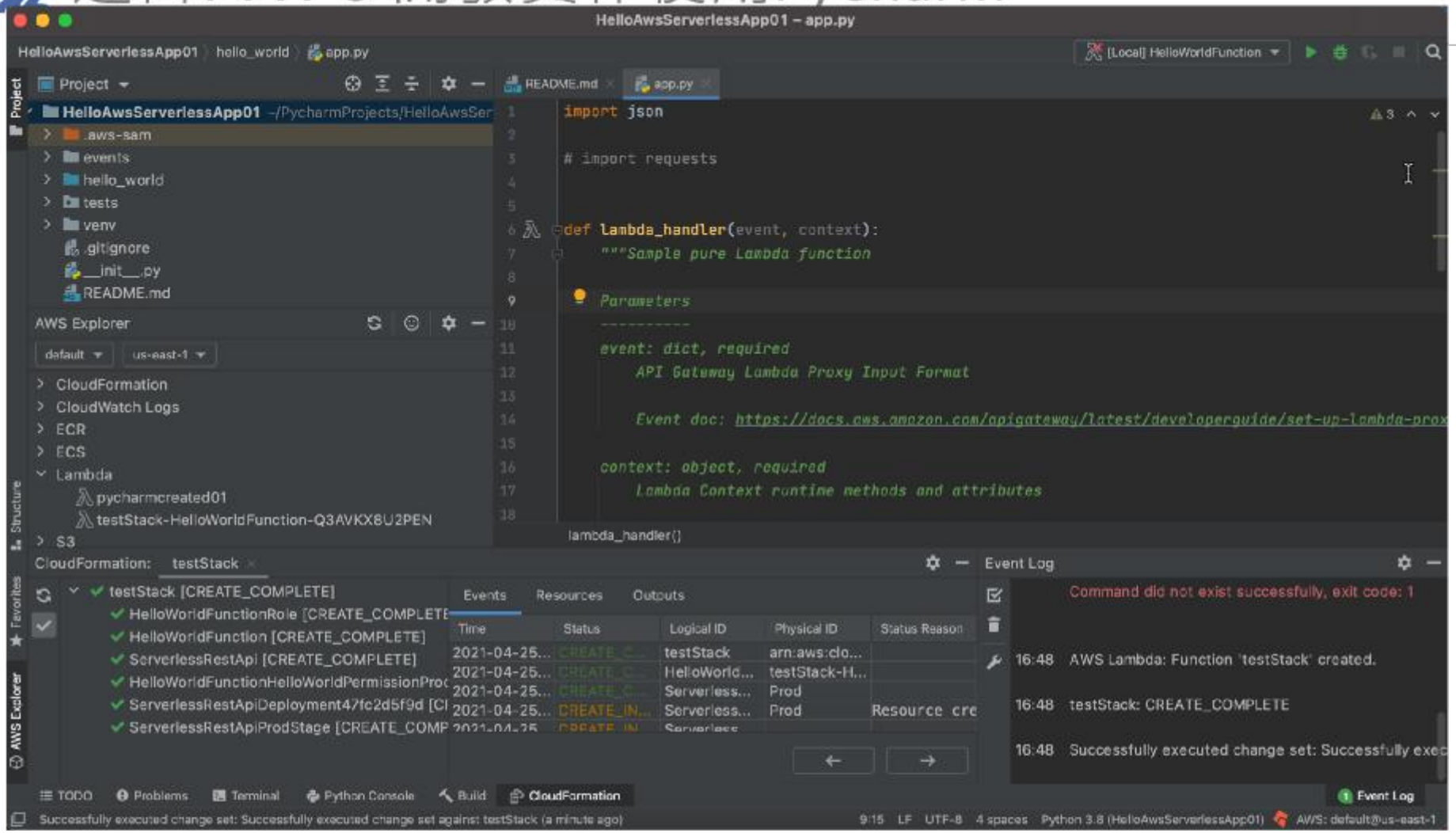
brew tap aws/tap
brew install aws-sam-cli <--installfailed

There seems to be an issue with the installer: aws/homebrew-tap#84
As a work-around you can force the install of the previous version (0.48.0):

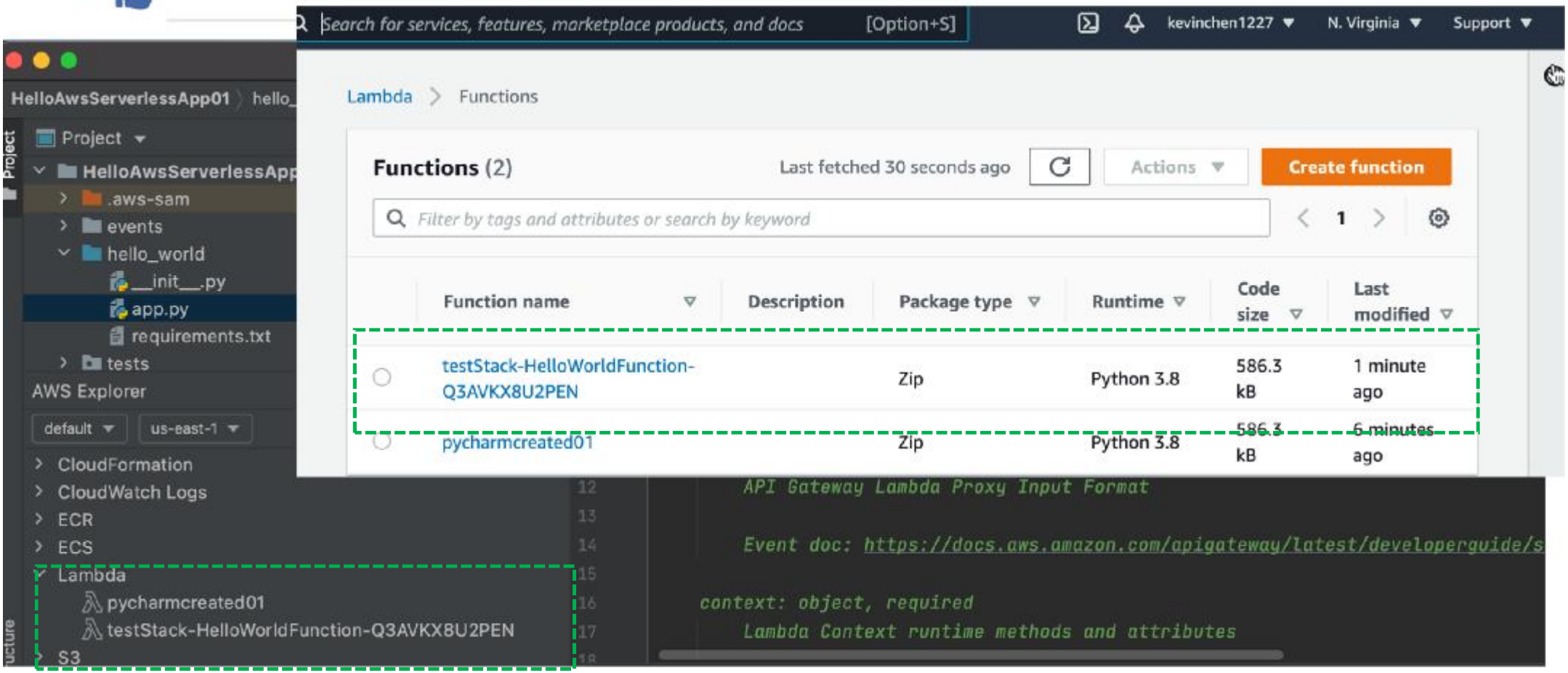
cd /usr/local/Homebrew/Library/Taps/aws/homebrew-tap
git reset --hard 9ddfdddbe1fed52940f9561ed04618682d5b24b2
HOMEBREW_NO_AUTO_UPDATE=1 brew install aws-sam-cli

Brew install aws-sam-cli
Brew upgrade aws/tap/aws-sam-cli

近日ＡＷＳ開發實作使用PyCharm



近日ＡＷＳ開發實作使用PyCharm



近日研讀ＡＷＳ Elastic Beantalk API 開發應用程式

最近程式設計-題目

1) The Spam Filter

When crawling the internet, we often encounter meaningless spam pages. A typical spamming technique is keyword stuffing, which stuffs a page with popular keywords, like "mp3" or "ipod," to increase its ranking in search engine results. We want to filter out such pages. Keyword-stuffed pages usually contain lots of machine-generated content, thus have less proper English sentences. Theoretically we may analyze the grammatical and semantic correctness for each sentence by natural language processing, but this would be computationally expensive. A lightweight alternative is to use a statistical technique, looking for probabilistic local consistency. We segment each document to n-grams of n consecutive words, where n is a small number such as 2, 3 … We define the frequency of the n-gram $w_{i+1} \dots w_{i+n}$ starting at word $i+1$ to be:

$$P(w_{i+1} \dots w_{i+n}) = \text{number of occurrences of the } n\text{-gram}$$

Note that n-grams are overlapping. For example, the third word of a document is covered by the first, second and third tri-gram, if the document has at least five words. Although they are overlapping, to simplify the computation we assume each of them is chosen independently to each other. We then define the probability of a document with k n-grams (and hence $k+n-1$ words) to be the product of the individual n-gram frequency, normalized by taking its k-th root:

$$\sqrt[k]{\prod_{i=0}^{k-1} P(w_{i+1} \cdots w_{i+n})}$$

Example

We are computing the bi-gram probability for a document consisting of a single sentence:

Don't cry because it is over, smile because it happened.

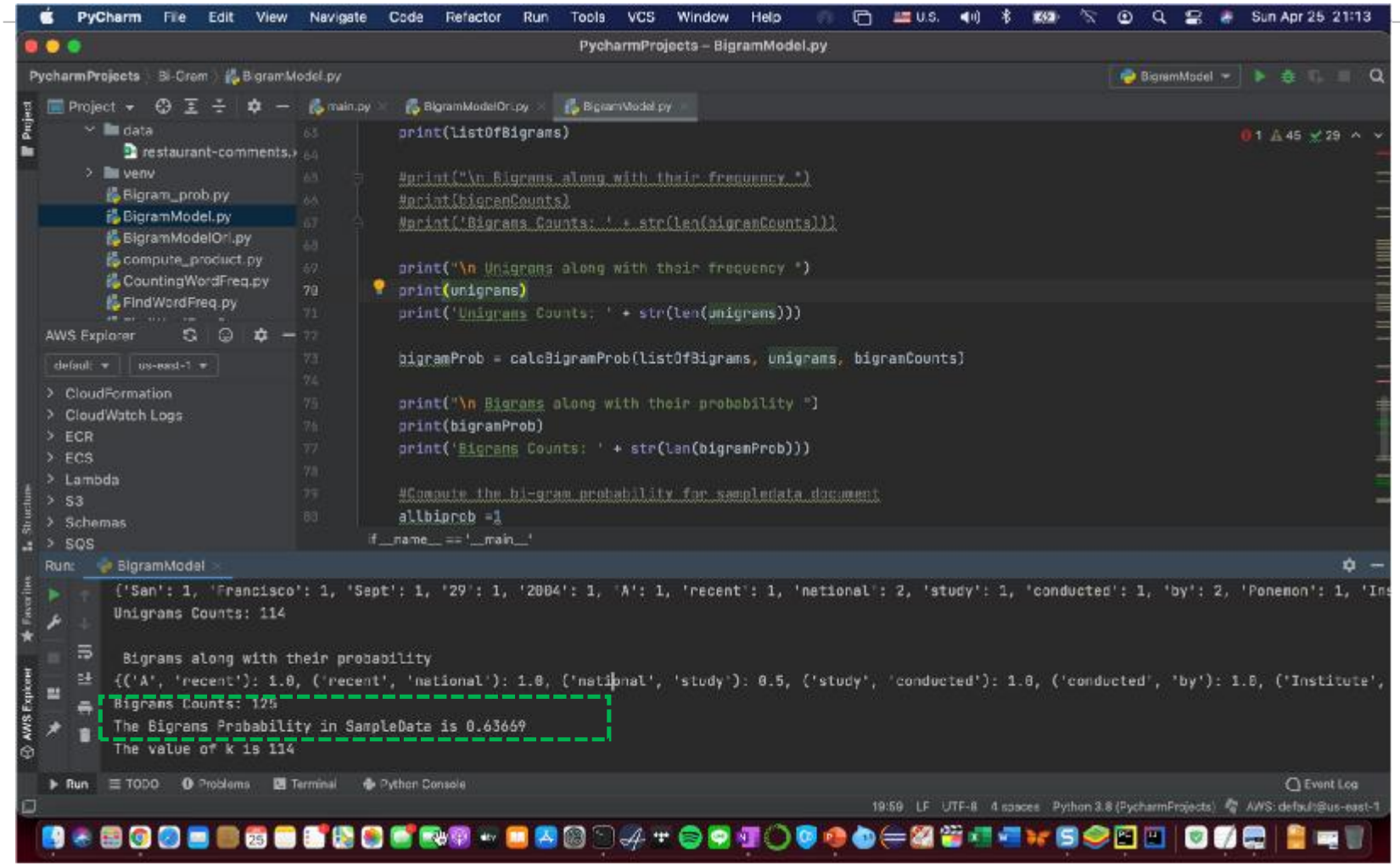
By the definition above, n equals 2 and k equals 9. The bi-gram probability can be calculated by:

(P(don't cry) * P(cry because) * P(because it) * P(it is) * P(is over) * P(over smile) *
P(smile because) * P(because it) * P(it happened)) ^ (1/9)
= (1 * 1 * 2 * 1 * 1 * 1 * 1 * 2 * 1) ^ (1/9)
= 4 ^ (1/9)
= 1.1665290395761165

Requirements

As an initial step of the spam filter project, you will write some code to read sample data from the attached text file, and calculate its *bi-gram probability*. Let's assume (1) words are case insensitive, and (2) words consist of letters and digits only. Feel free to read reference manuals or search the internet.

最近程式設計-答案


print(listOfBigrams)
#print("\n Bigrams along with their frequency ")
#print(bigramCounts)
#print('Bigrams Counts: ' + str(len(bigramCounts)))
print("\n Unigrams along with their frequency ")
print(unigrams)
print('Unigrams Counts: ' + str(len(unigrams)))
bigramProb = calcBigramProb(listOfBigrams, unigrams, bigramCounts)
print("\n Bigrams along with their probability ")
print(bigramProb)
print('Bigrams Counts: ' + str(len(bigramProb)))
#Compute the bi-gram probability for sampledata document
allbiprob =1
if __name__ == '__main__'
{'San': 1, 'Francisco': 1, 'Sept': 1, '29': 1, '2004': 1, 'A': 1, 'recent': 1, 'national': 2, 'study': 1, 'conducted': 1, 'by': 2, 'Ponemon': 1,
Unigrams Counts: 114
Bigrams along with their probability
{('A', 'recent'): 1.0, ('recent', 'national'): 1.0, ('national', 'study'): 0.5, ('study', 'conducted'): 1.0, ('conducted', 'by'): 1.0, ('Institute',
Bigrams Counts: 125
The Bigrams Probability in SampleData is 0.63669
The value of k is 114

最近程式設計-題目

Question

The TF-IDF algorithm is a classical algorithm used to find the most important words in a set of documents.

In this exercise, we only focus on the TF (term-frequency) part of this algorithm: you have to compute the frequency of words in a sentence.

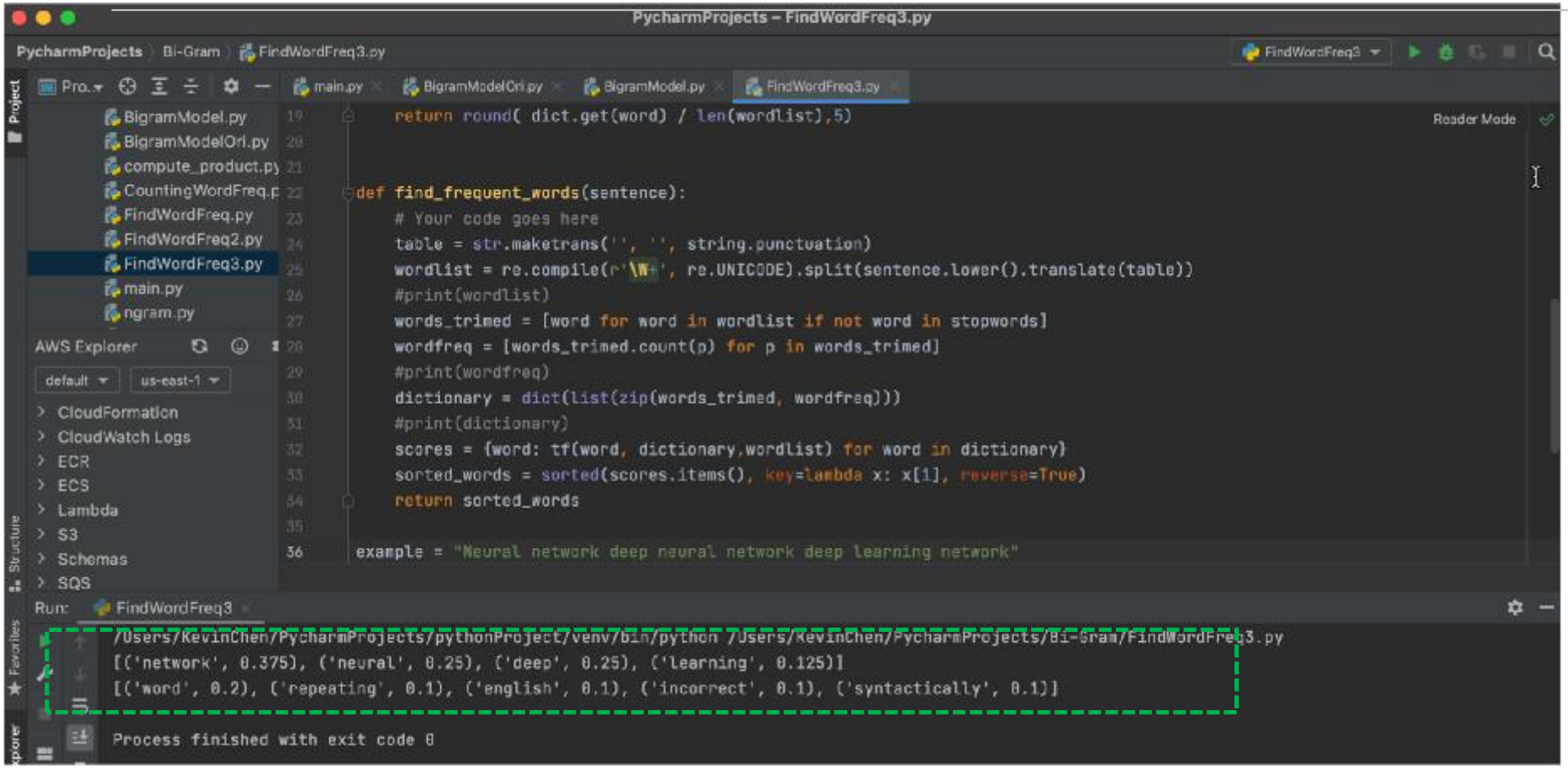
The ***find_frequent_words()*** function must return a list of tuples giving the frequency of words from the ***sentence*** parameter.

Output example:
[('deep', 0.25), ('learning', 0.125), ('network', 0.375), ('neural', 0.25)]

The order of tuples is not significant.
All sentences are in English.
The ***stop_words*** package is available should you need it.

As a Data Scientist, we expect you to clean the data in a way you find the most appropriate.

最近程式設計-答案


PycharmProjects – FindWordFreq3.py
PycharmProjects › Bi-Gram › FindWordFreq3.py
main.py
BigramModelOri.py
BigramModel.py
FindWordFreq3.py
Reader Mode
BigramModel.py
BigramModelOri.py
compute_product.py
CountingWordFreq.p
FindWordFreq.py
FindWordFreq2.py
FindWordFreq3.py
main.py
ngram.py
AWS Explorer
default
us-east-1
CloudFormation
CloudWatch Logs
ECR
ECS
Lambda
S3
Schemas
SQS
return round(dict.get(word) / len(wordlist),5)
def find_frequent_words(sentence):
# Your code goes here
table = str.maketrans('', '', string.punctuation)
wordlist = re.compile(r'\W+', re.UNICODE).split(sentence.lower().translate(table))
#print(wordlist)
words_trimed = [word for word in wordlist if not word in stopwords]
wordfreq = [words_trimed.count(p) for p in words_trimed]
#print(wordfreq)
dictionary = dict(list(zip(words_trimed, wordfreq)))
#print(dictionary)
scores = {word: tf(word, dictionary,wordlist) for word in dictionary}
sorted_words = sorted(scores.items(), key=lambda x: x[1], reverse=True)
return sorted_words
example = "Neural network deep neural network deep learning network"
Run: FindWordFreq3
/Users/KevinChen/PycharmProjects/pythonProject/venv/bin/python /Users/KevinChen/PycharmProjects/Bi-Gram/FindWordFreq3.py
[('network', 0.375), ('neural', 0.25), ('deep', 0.25), ('learning', 0.125)]
[('word', 0.2), ('repeating', 0.1), ('english', 0.1), ('incorrect', 0.1), ('syntactically', 0.1)]
Process finished with exit code 0

最近程式設計-題目

2) SQL

Design database tables to support following requirements -

1. Each user has an unique email and a user name.

2. Each user belongs to a company that has a company name.

3. Each user is assigned with a certain application role - admin, analyst or viewer.

4. Users can create lots of dashboards for analysis purpose, and each dashboard contains a dashboard name and a json config string.

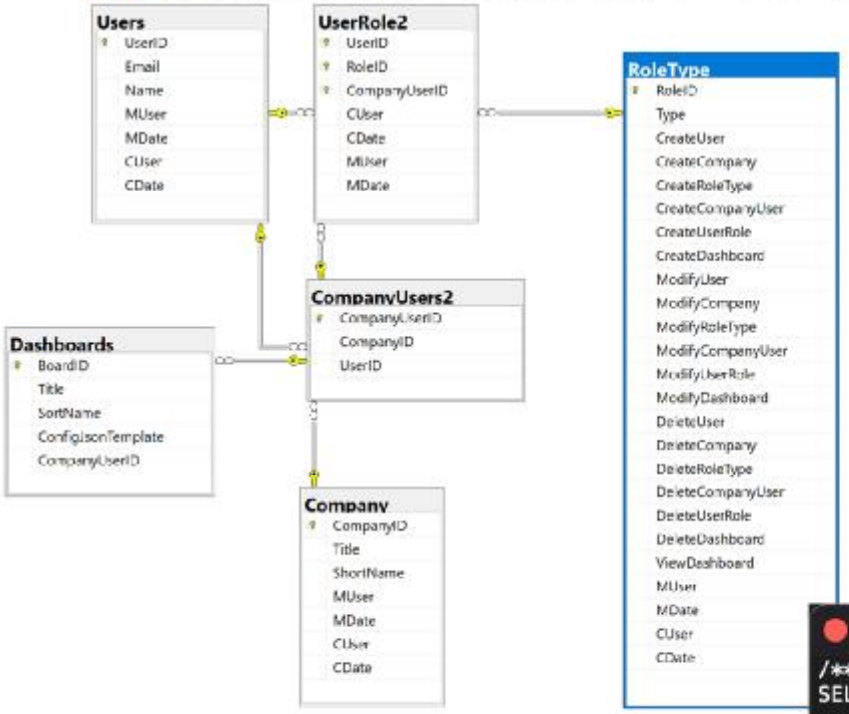
You can design tables using any tools and take screenshots, or on papers and take photos.

Based on above design, please provide a SQL query to list how many dashboards owned by admin for each company, sorted by company name.

Requirements

Result contains (1) an image file with your table design and (2) a text file with your SQL query.

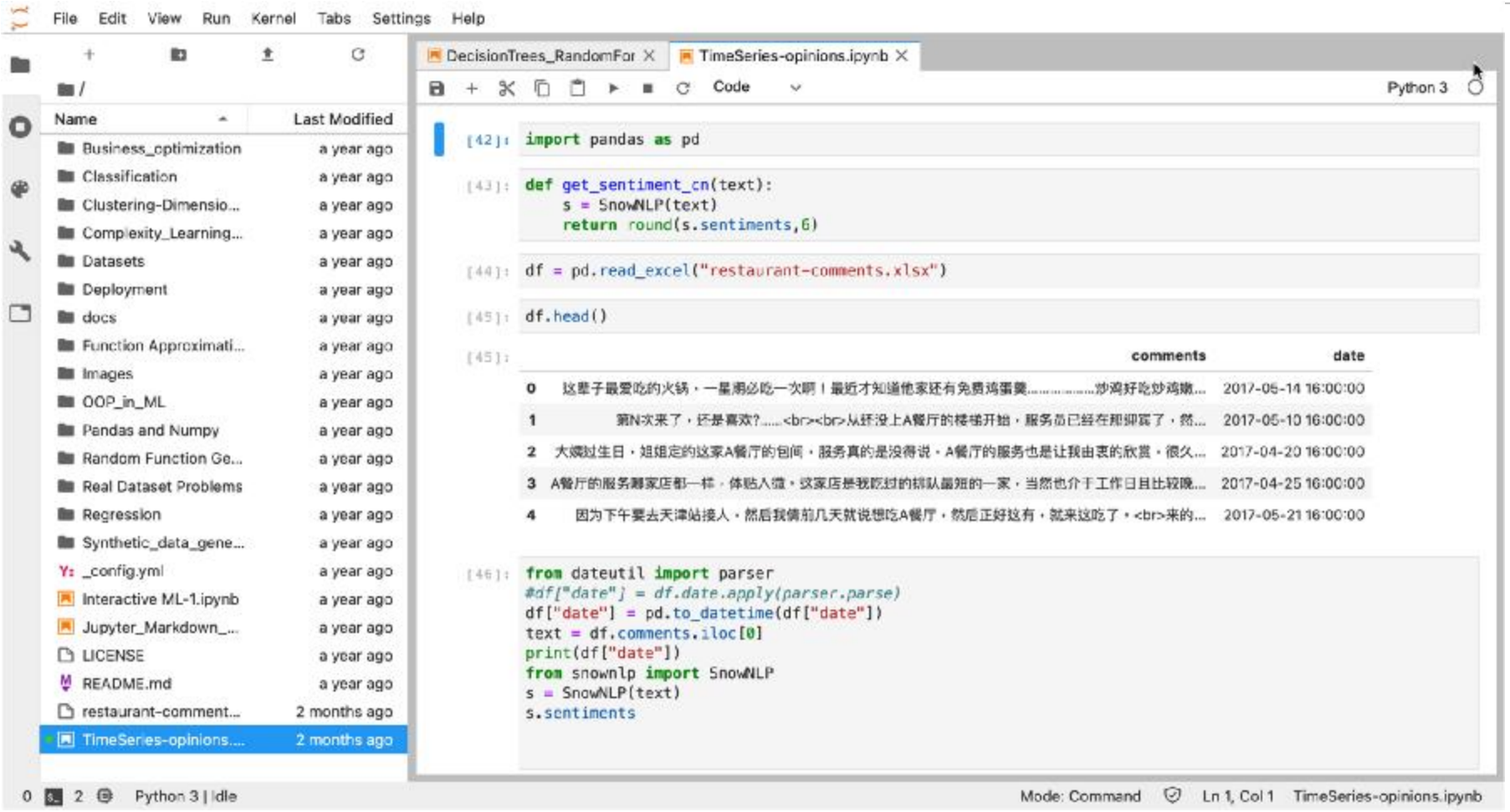
最近程式設計-答案



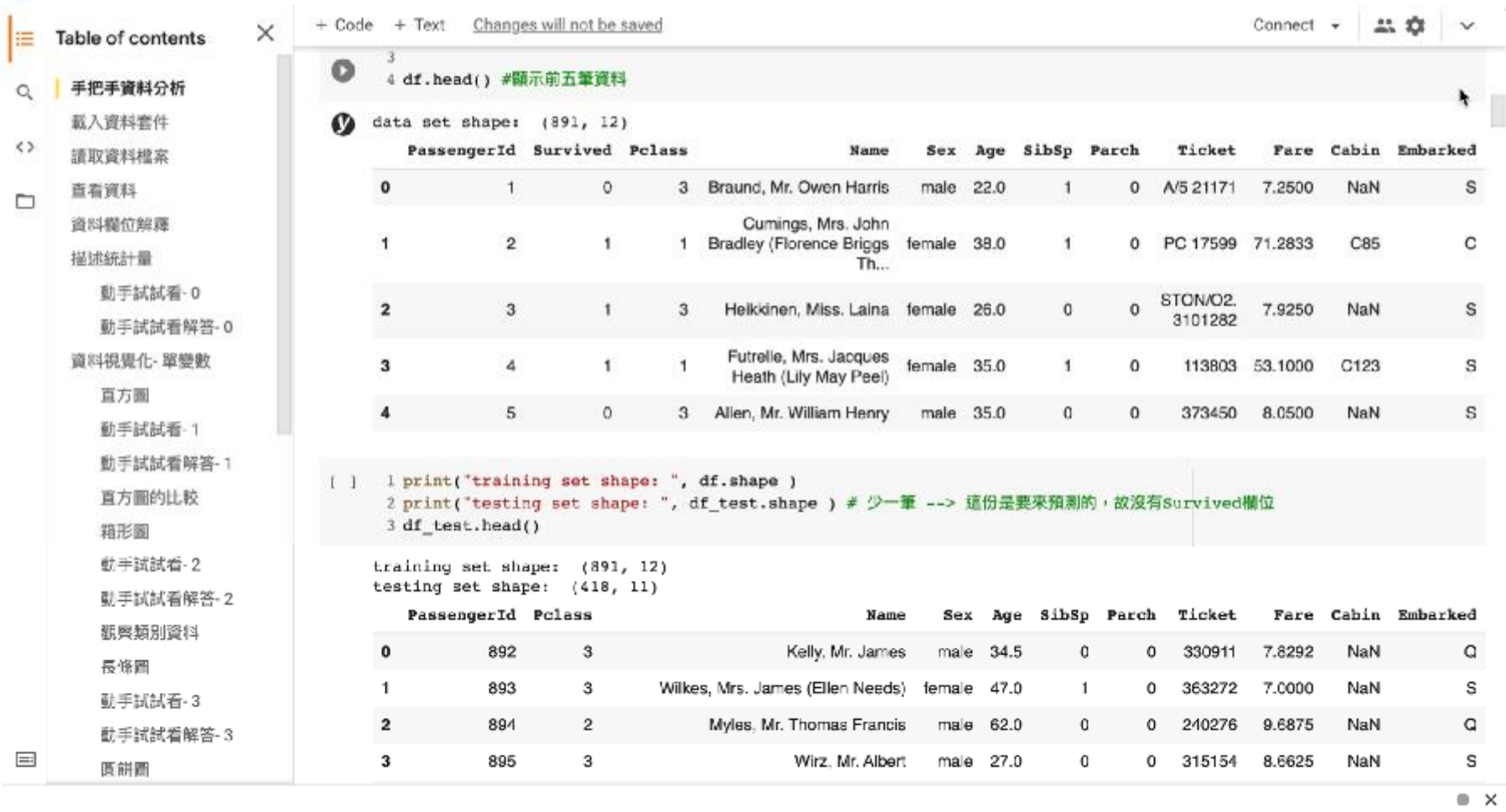
(2) SQLQuery.sql

```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT [board].[boardID]
      ,[board].[Title]
      ,[board].[SortName]
      ,[board].[ConfigJsonTemplate]
            ,[Rtype].[Type]
            , [Company].ShortName as Company
  FROM ((([NetBase].[dbo].[Dashboards] [board] INNER JOIN [CompanyUsers2] [comuser]
      ON [board].CompanyUserID = [comuser].[CompanyUserID])
           INNER JOIN [Company] on [Comuser].CompanyID = [Company].CompanyID)
           INNER JOIN [UserRole2] [Role] ON [comuser].[UserID]= [Role].UserID)
           INNER JOIN [RoleType] [Rtype] on [Role].RoleID = [Rtype].RoleID
        WHERE [Rtype].[Type]='admin'
        ORDER BY [Company].ShortName ASC
```

最近AI程式設計-網友意見分析



最近AI程式設計-鐵達尼號死亡乘客分析



	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN	S
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85	C
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN	S
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123	S
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN	S

	PassengerId	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin	Embarked
0	892	3	Kelly, Mr. James	male	34.5	0	0	330911	7.8292	NaN	Q
1	893	3	Wilkes, Mrs. James (Ellen Needs)	female	47.0	1	0	363272	7.0000	NaN	S
2	894	2	Myles, Mr. Thomas Francis	male	62.0	0	0	240276	9.6875	NaN	Q
3	895	3	Wirz, Mr. Albert	male	27.0	0	0	315154	8.6625	NaN	S

Azure Cloud/Office 365技術經驗

1. 高雄市衛生局智慧心理健康雲規畫

PM：由RFP到提案

SA/Architect ：採用Azure平台之Azure AD,Web App管理使用者帳號與權限、SQL Service整合各業務單位個案資料資歷、 AzureML與演算法提供智慧化與自動化機制於效能管理與個案管理與宣導排程，Azure LUIS與QnAMaker設計無人客服。

技術實作測試Web App, AzureML, AzureLUIS與QnAMaker以及串接LINE app

規畫心理健康管理APP：規劃遊戲機制、與健康存摺。

焦點專案

高雄市衛生局智慧心理健康雲端資料庫-目標與效益

心理健康管理APP

- 提供民眾自我心理檢測
- 主動提供生活處方與資訊
- 主動推播健康/衛教資訊
- 遊戲任務提高使用黏著度

社區民眾端

智慧心理雲端資料庫

- 大數據分析資料主動產出
- 掌握各類服務個案
- 以利後續政府政策擬定

個管人員端

心衛中心決策端

個管人員APP

- 提供即時便利的工具
- 提升訪視品質
- E化取代紙本通報並即時派案
- 整合簡化作業流程

心衛中心管理端

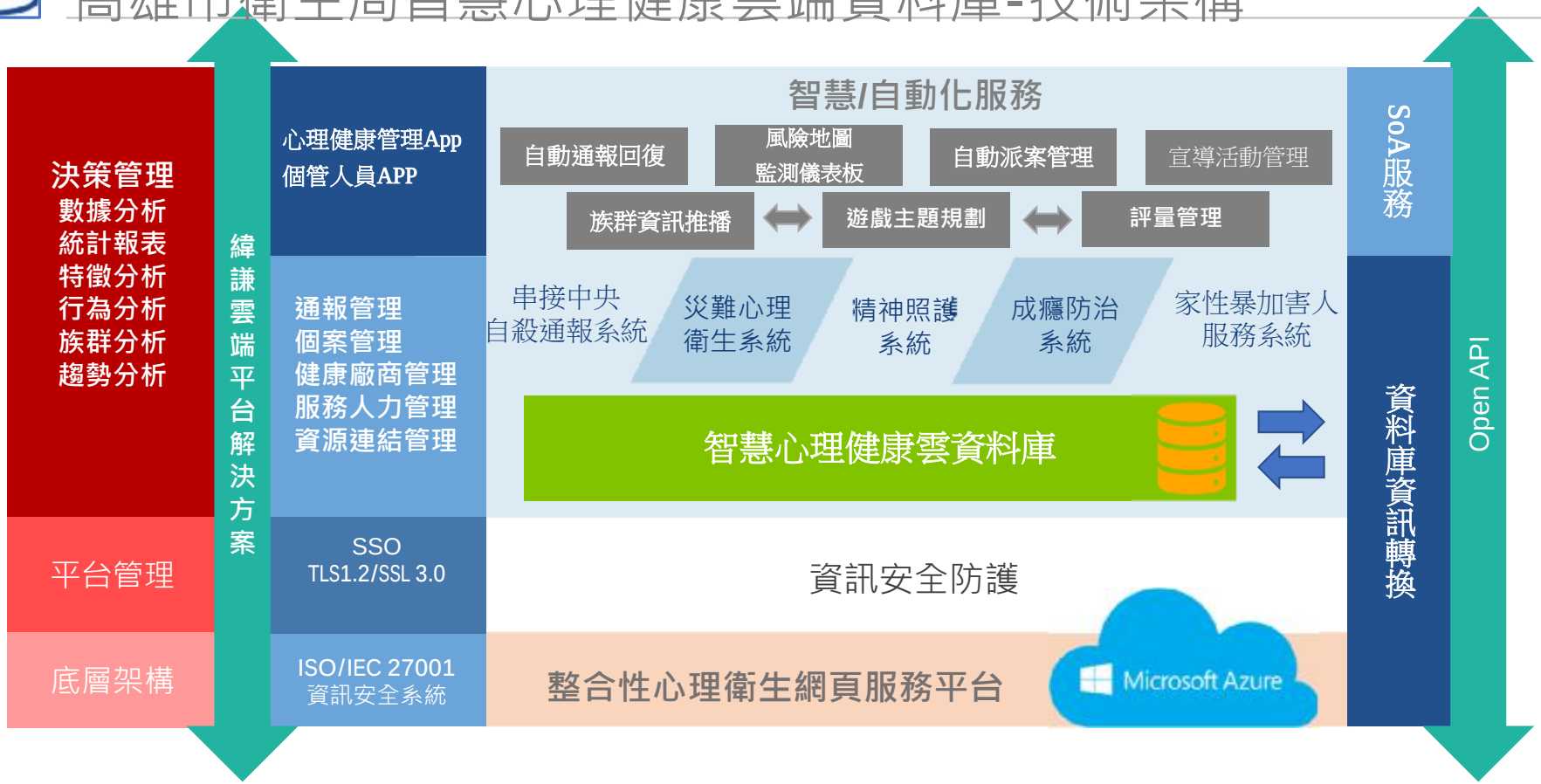
整合性心理衛生網頁服務平台

- 快速掌握個案有助於評估管理
- 監控處理服務時效，提昇服務品質
- 掌握各項統計數據精進服務策略

18

焦點專案

高雄市衛生局智慧心理健康雲端資料庫-技術架構



Azure Cloud/Office 365技術經驗

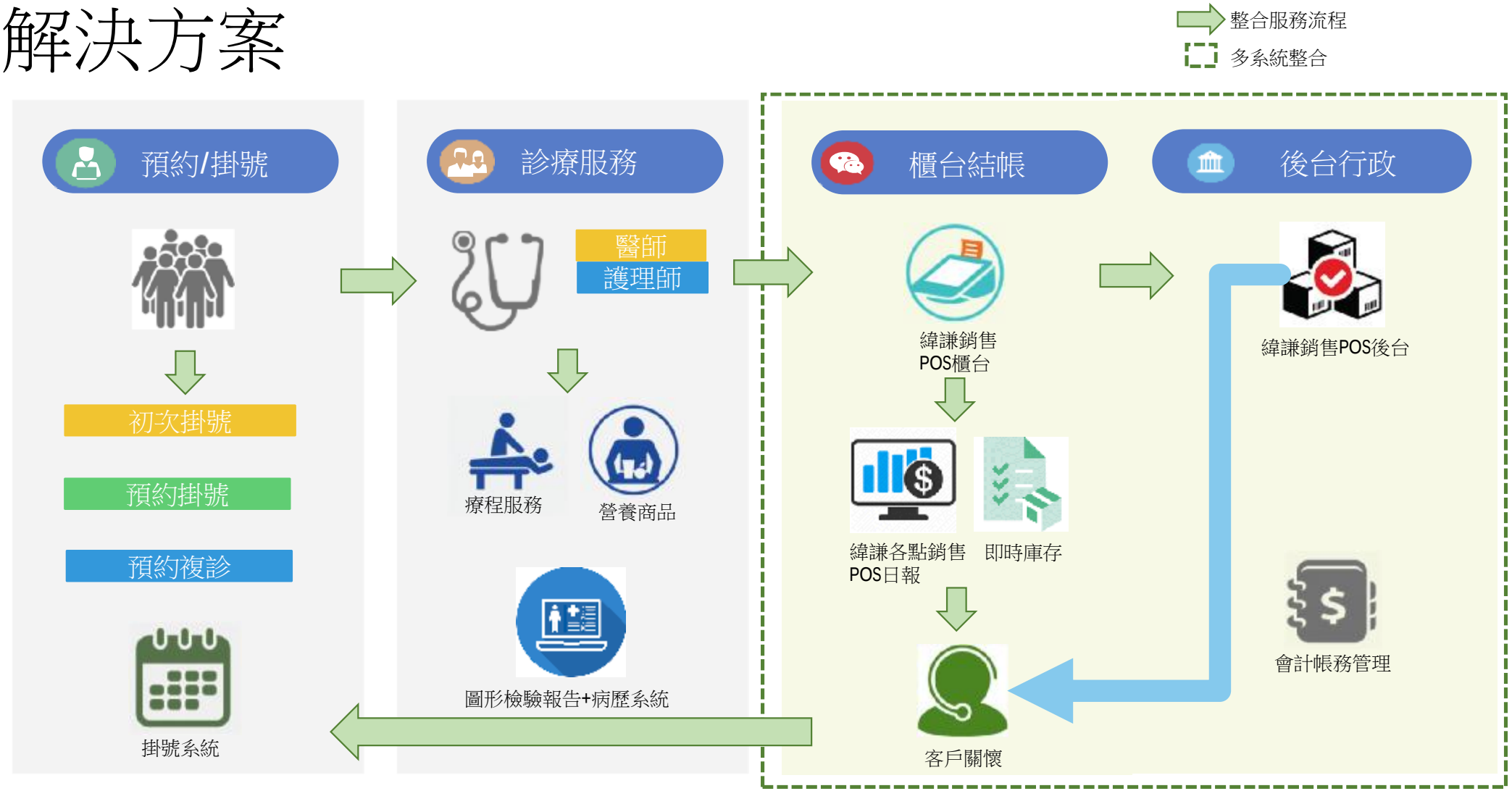
2.菁英診所健檢雲端系統

PM：由Engagement到專案開發

SA/Architect ：採用Azure Ubuntu VM, SQL service、Azure ML推薦客戶診條服務與健康食品、建議客服主動健康關懷時程、銷售預測

DBA: 設計電子病歷的資料庫有關進銷存以及預約報到管理

解決方案


整合服務流程
多系統整合
預約/掛號
診療服務
櫃台結帳
後台行政
初次掛號
預約掛號
預約複診
掛號系統
醫師
護理師
療程服務
營養商品
圖形檢驗報告+病歷系統
緯謙銷售
POS櫃台
緯謙各點銷售
POS日報
即時庫存
客戶關懷
緯謙銷售POS後台
會計帳務管理

Azure Cloud/Office 365技術經驗

南亞塑膠銅箔基板廠提升含浸動用率AI專案

- **目標與效益**

  提供生產產品處方參數建議，以減少生產產品規格換製時間，提升可生產時間及達成動用率提升。

- **專案角色：PM, SA, SD,**
  - **DBA, Python PG：Data Injection部分**
- **技術架構**
  - 採用微軟Azure雲端平台之AzureML, MLOps, SQL Service, Docker封裝與IoT Edge邊緣運算，以及PowerBI

    VS.net + SSIS + SSRS: 收集生產與品質紀錄，依據規則以程式初步資料清洗轉成CSV檔並備份。

    Python + Azure ML: 以Python於AzureML處理生產紀錄萃取特徵、訓練AI模型。

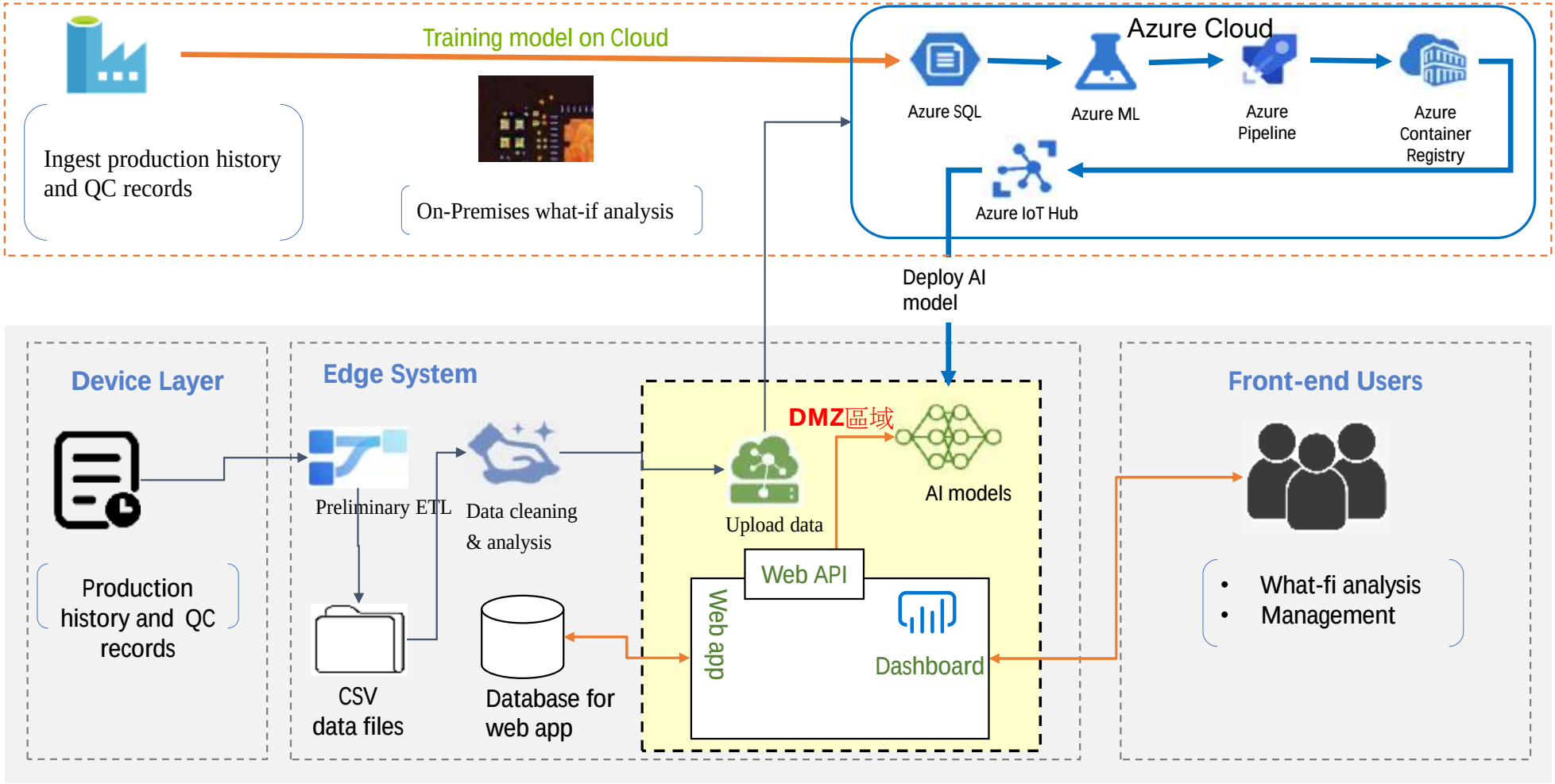
    Azure MLOps: 發展與管理資料建模流程版本從資料、程式到模型，並自動化部屬模型於邊緣運算做即時監測。

    可視化儀表板: 開發生產模擬分析（what-if分析）透過AI模型的restFul api與前端介面溝通。

    邊緣運算平台：Python Flask以及DASH作為應用系統平台

Azure Cloud/Office 365技術經驗

南亞塑膠銅箔基板廠提升含浸動用率AI專案-技術架構



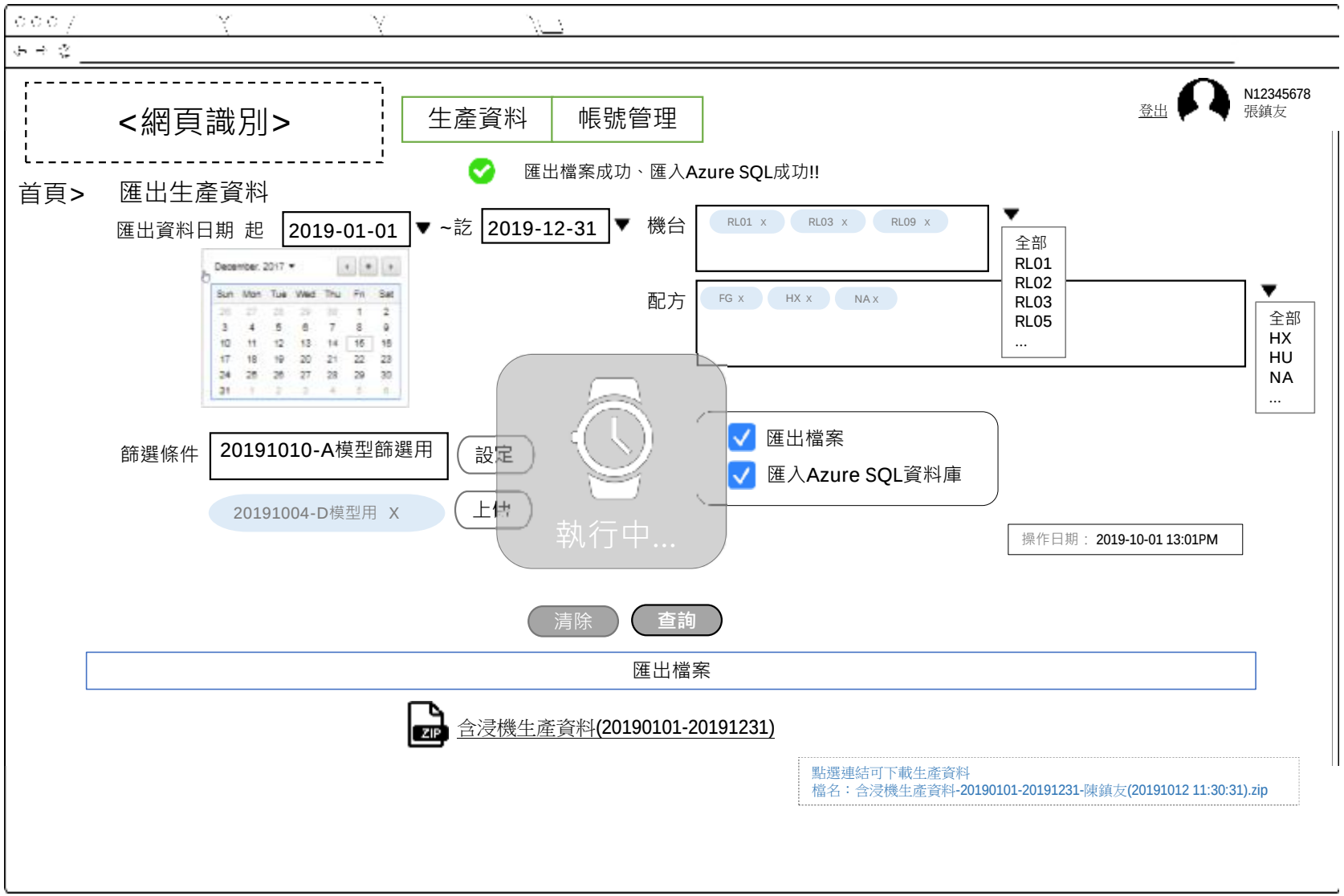
畫面需求
-3. 生產資料

3.1匯出生產資料查詢

3.3 匯入資料篩選條件

3.4 匯入Azure SQL資料庫

系統紀錄使用者操作紀錄



Azure Cloud/Office 365技術經驗

緯創數位轉型導入緯穎i4.0 Armstrong on Azure for Wiwynn

- **目標與效益**

  瞄準於建構一個行動戰情室，顯示緯創全球所有工廠各車間所有生產機台的生產狀況，即時行動化緊急通知業管人員能快速解決問題、提供總公司透明化生產與品質監控。

- **技術架構**

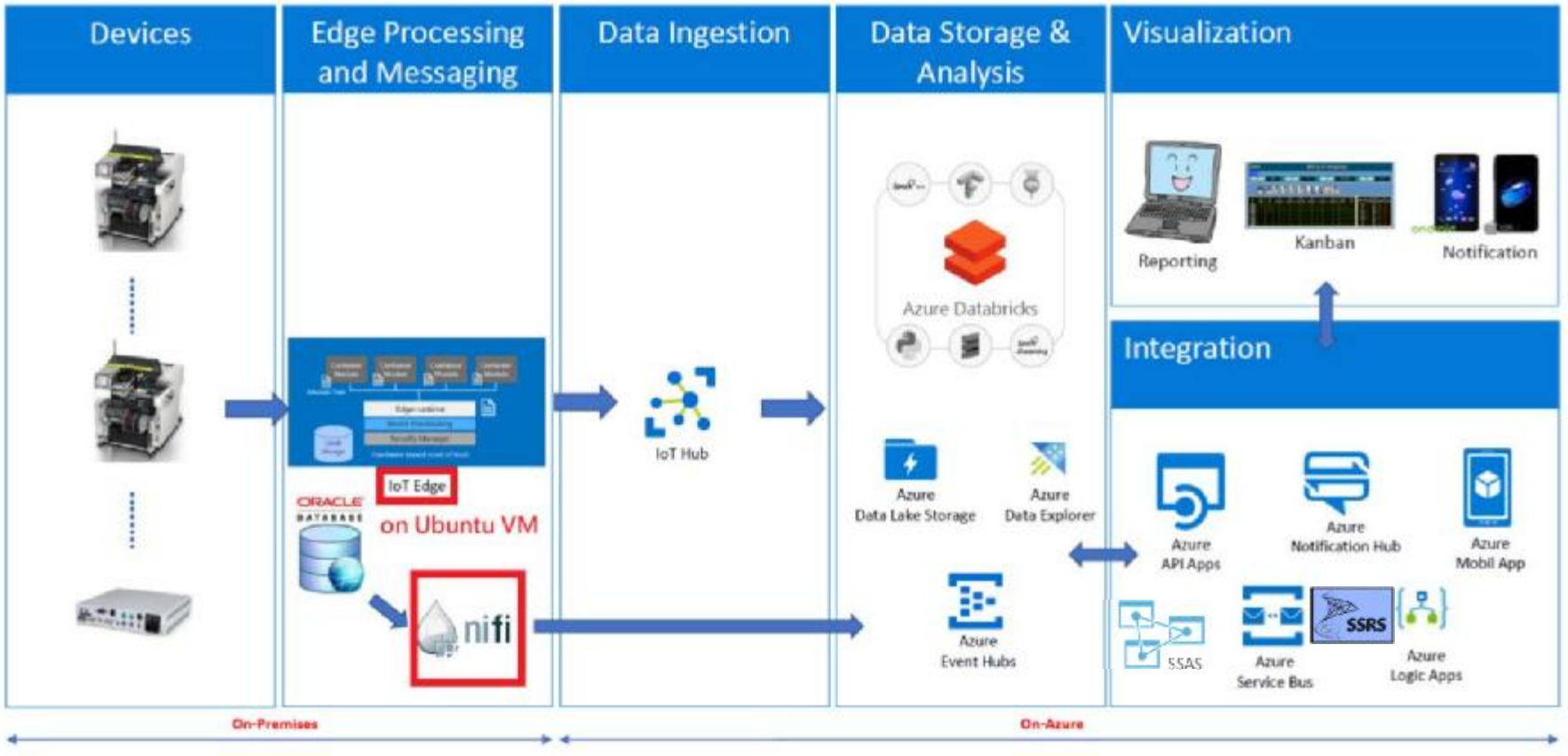
  - **Data Streaming技術，採用微軟Azure雲端個多種服務，（詳如下圖）**

    Apache Kafka+NiFi:採集生產記錄，存放於邊緣運算的資料庫，將資料上傳到Azure雲端儲，進而分析產生有意義觀測指標顯示於儀錶板。

VS.net + SSIS + SSRS + SSAS: 設計ETL以及靜態報表、多維度報表分析。

- **專案角色：Data Lead in vendor side, 帶領member開發AngularJS**

緯創數位轉型導入緯穎i4.0 Armstrong on Azure for Wiwynn-技術架構



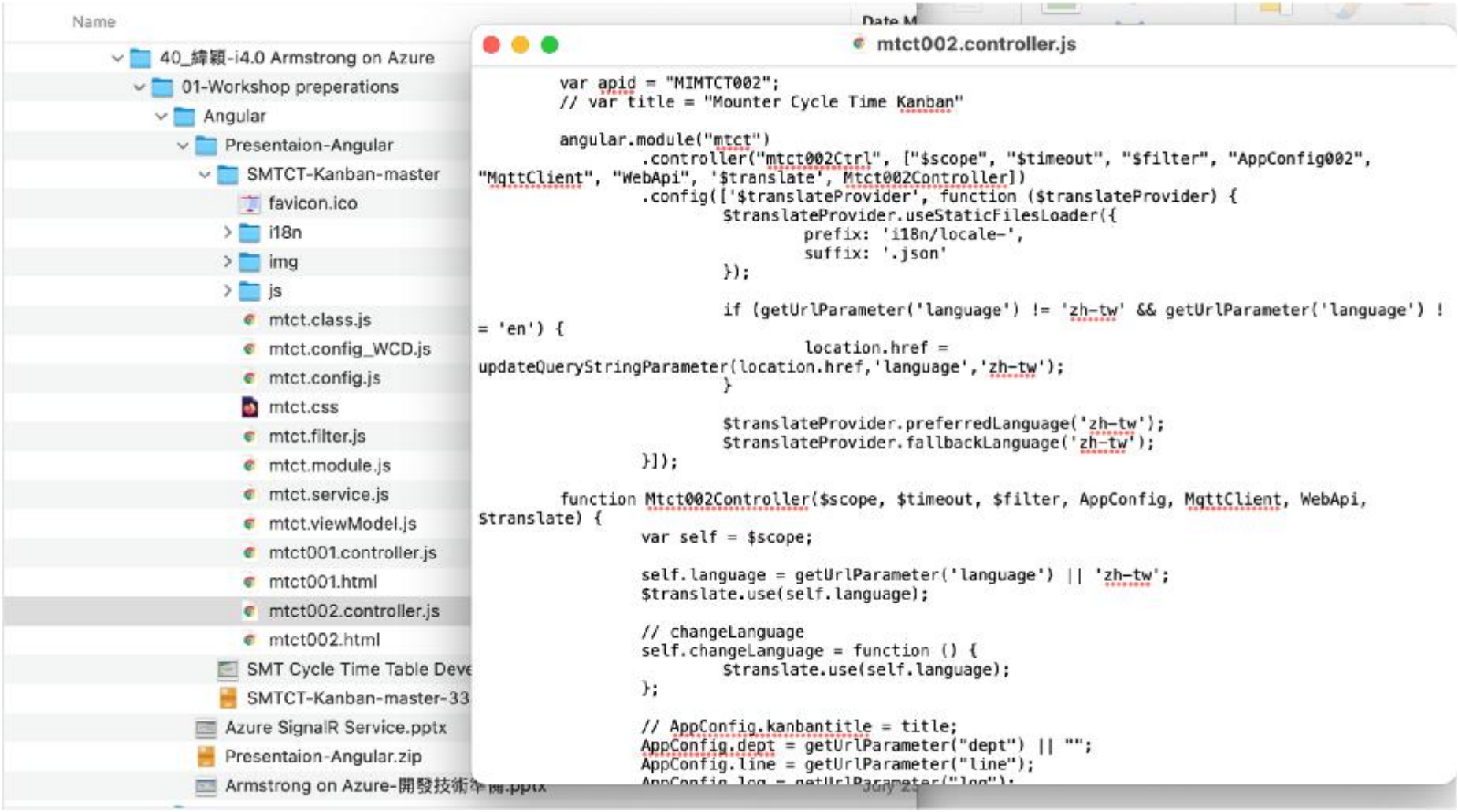
SMT Cycle Time Table 開發

項目	定義	數據來源	更新頻率&判定標準
1	機台序號	依產線實際Layout, 由第一台設備开始排序	/
2	機台名稱	依實際Layout顯示機台類型	/
3	實時Cycle Time	Mounter：Pana CT文件 其他設備：KVM技術	/
4	Bottle Neck: 瓶頸站標準CT	MIC系統，由PSE Maintain	/
5	Standard Balance Rate：	公式：各機台標準CT相加 / 瓶頸站標準CT*機台數量 各機台標準CT由MIC系統取得	/
6	Actual Balance Rate：	計算公式：各機台“CT平均值”加 總/ “CT平均值”最大值*機台數量 1. 各機台根據30分鐘內生產的板數取“CT平均值” 2. 若機台有拋出CT文件則參與Actual Balance Rate的計算，未拋出CT文件則不參與計算並顯示NA	更新頻率： 每5分鐘更新顯示前30分鐘的平均Balance Rate（例, 9:00 顯示8:30~9:00的數值, 9:05顯示8:35~9:05的數值) 判定標準： 大於或等於Standard Balance Rate顯示綠色，小於 显示红色 異常通知：詳見Page 5

線別	M1 Printer	M2 SPI	M3 Mounter1	M4 Mounter2	M5 Mounter3	M6 Mounter4	M7 Mounter5	M8 Mounter6	M9 Reflow	M10 AOI	Bottle Neck	Standard Balance Rate	Actual Balance Rate
CS2A	14.8	13.0	24.0	15.8	15.2	15.1	8.5	7.5	11.5	12.3	16	92%	57%
CS2B	14.8	12.0	13.3	11.7	14.3	22.0	13.2	12.8	12.2	13.7	15	94%	64%
CS3A	16.2	14.1	19.2	18.7	17.4	16.9	18.6	14.3	11.7	15.6	20	97%	91%
CS3B	17.1	14.3	17.2	17.8	18.6	17.5	13.7	15.1	12.0	15.8	18	96%	92%
CS4A	16.3	15.0	9.8	10.1	9.9	9.7	10.1	10.8	11.2	14.8	17	93%	72%
CS4B	11.2	13.8	12.3	12.1	13.8	10.7	13.6	13.7	10.7	11.3	14	92%	91%
CS5A	12.8	14.0	17.3	17.6	17.1	17.4	13.9	14.2	12.8	14.2	18	94%	93%
CS5B	11.7	15.1	14.7	13.8	11.6	12.7	15.8	11.5	11.0	15.2	16	95%	84%

注：1. 首列的設備名稱以機台數量最多的線別為準，沒有此機台的線別，其C/T欄位顯示NA
2. 中文字體：
微軟正黑體
英文字體：
Arial (本文)

AngularJS開發 SMT Cycle Time Table



Data Streaming Using Apache Kafka開發經驗

Go to Docker official web site & base on O.S to download and install “Docker Community Edition”

- https://store.docker.com/search?type=edition&offering=community

Verify Docker readiness
docker --version

docker-compose --version

docker run hello-world

Create Training File/Directory (linux & mac only)

mkdir ~/datafabric

mkdir ~/datafabric/01_software

mkdir ~/datafabric/02_document

mkdir ~/datafabric/03_workspace

mkdir ~/datafabric/03_workspace/env

mkdir ~/datafabric/03_workspace/hands-on

Get Docker-Compose configuration file

wget
https://gist.githubusercontent.com/erhwenkuo/7b72c2464419ab5806a7332005ae41e8/raw/c4b1710fc2b77ce7b90adfee6646cd93fb349cbd/docker-compose.yml

mv docker-compose.yml ~/datafabric/03_workspace/env

Setup Apache Kafka & ZooKeeper

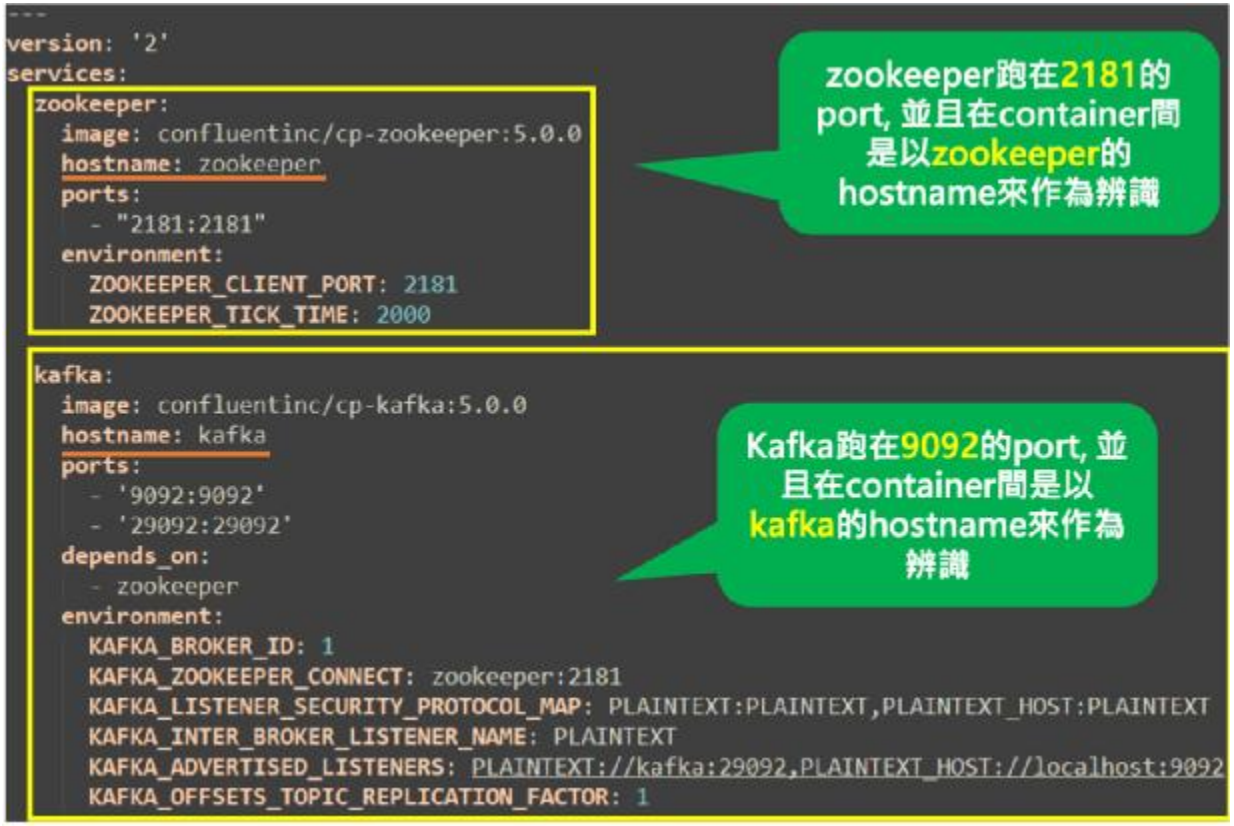
```
# Change to path that contains docker-compose.yml  (linux & mac only)

cd ~/datafabric/03_workspace/env

# Start Zookeeper & Kafka

docker-compose up -d

# Verify Zookeeper & Kafka services

docker-compose ps

# Verify Zookeeper is healthy (linux & mac only)

docker-compose logs zookeeper | grep  -i binding

# Verify Kafka is healthy (linux & mac only)

docker-compose logs kafka | grep  -i started


# Get into Docker container

docker exec -it env_kafka_1 bash
```

```
# Create a topic

kafka-topics --create \

  --topic test \

  --replication-factor 1 \

  --partitions 1 \

  --zookeeper zookeeper:2181

# Publish data

kafka-console-producer \

  --broker-list kafka:9092 \

  --topic test

# Subscribe data

kafka-console-consumer \

  --bootstrap-server kafka:9092 \

  --topic test \

  --from-beginning
```

```
# Shutdown Zookeeper & Kafka

docker-compose stop

# Start exiting Zookeeper & Kafka

docker-compose start

# Remove Zookeeper & Kafka container/data

docker-compose down
```

Data Streaming Using Apache Kafka開發經驗

```
---
version: '2'
services:
  zookeeper:
    image: confluentinc/cp-zookeeper:5.0.0
    hostname: zookeeper
    ports:
      - "2181:2181"
    environment:
      ZOOKEEPER_CLIENT_PORT: 2181
      ZOOKEEPER_TICK_TIME: 2000

  kafka:
    image: confluentinc/cp-kafka:5.0.0
    hostname: kafka
    ports:
      - '9092:9092'
      - '29092:29092'
    depends_on:
```

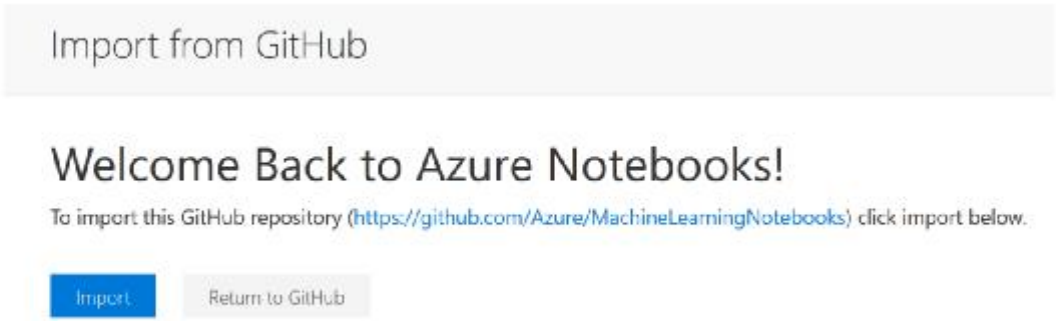


Data Streaming Using Apache Kafka開發經驗

```
public static void main(String[] args) throws Exception {
     // 步驟1. 取得Kafka的Producer實例
    final Producer<String, byte[]> producer = getKafkaProducer();

    // 步驟2: 設定要發佈的topic名稱
    String topicName = "ak04.hw." + STUDENT_ID;

    // 步驟3: 讀取讀取CSV的檔案(注意: nyc_taxi_data2.csv必需要在專案的根目錄
    String csv_file = "nyc_taxi_data2.csv";

    int validRecordCount = 0; // 用來計算有效的資料筆數
    int invalidRecordCount = 0; // 用來計算無效的資料筆數
    int rowNumber = 0; // 用來

    // 如何逐行讀取CSV的每一行資料
    try {
        File csv_data = new File(csv_file);
        BufferedReader csv_data_buffer = new BufferedReader(new FileReader(csv_data));
        String lineContent = "";

        // 迭代讀取每一行的資料
        while((lineContent = csv_data_buffer.readLine()) != null) {
            rowNumber++;

            if(rowNumber == 1)
                continue; // 由於第一行是header, 所以我們要ignore

            // 進行每一行的資料檢查或相對應的處理
            if(lineContent==null || lineContent.isEmpty())
                invalidRecordCount++;
            else {
                // 進行字串轉換解析成Avro物件
                Taxidata taxidata = parseCsvLineToAvro(rowNumber, lineContent);

                // 檢查taxidata是否為null
                if(taxidata==null)
                    invalidRecordCount++;
                else {
                    // 這是一個有效的資料行
                    validRecordCount++;

                     System.out.println(rowNumber-1 + "," + taxidata);
                    RecordMetadata recordMetadata = producer.send(new ProducerRecord<>(
                            topicName,
                            "" + (rowNumber - 1),
                            serializeToByte(taxidata))).get(); // 讓我們使用sync的手法來看每一筆的回應
```

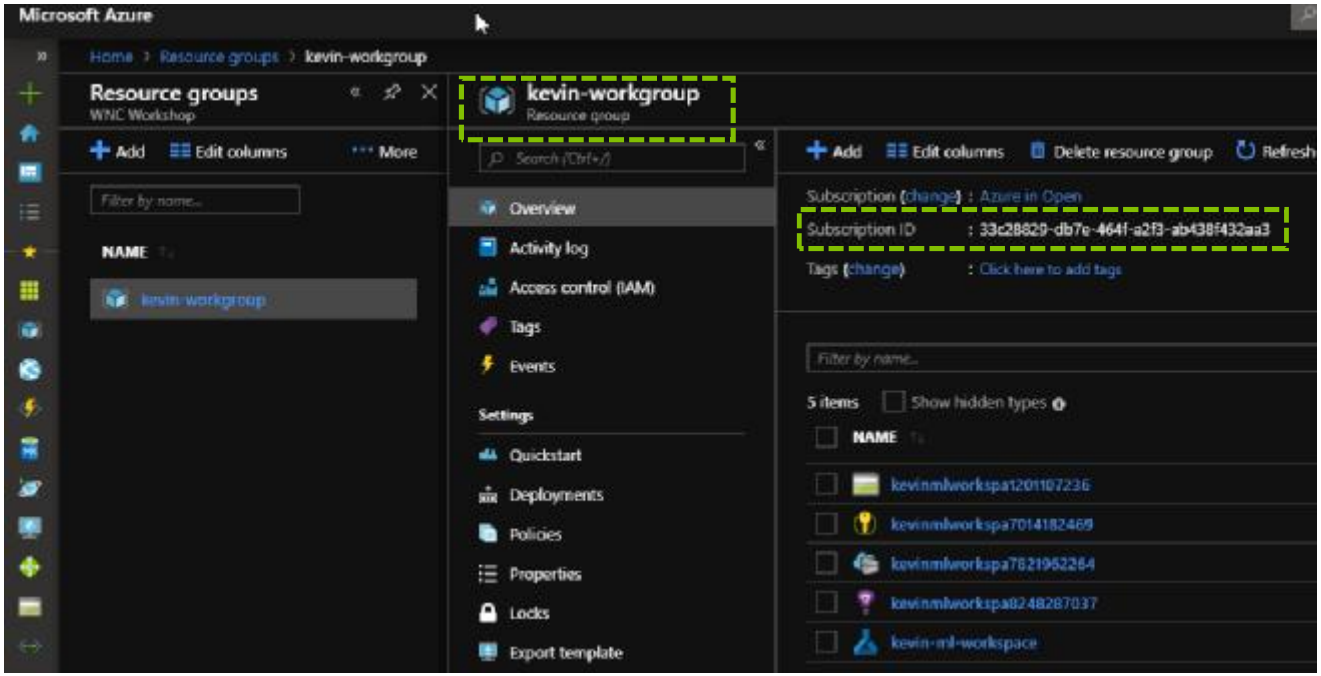
Launch Azure Notebook

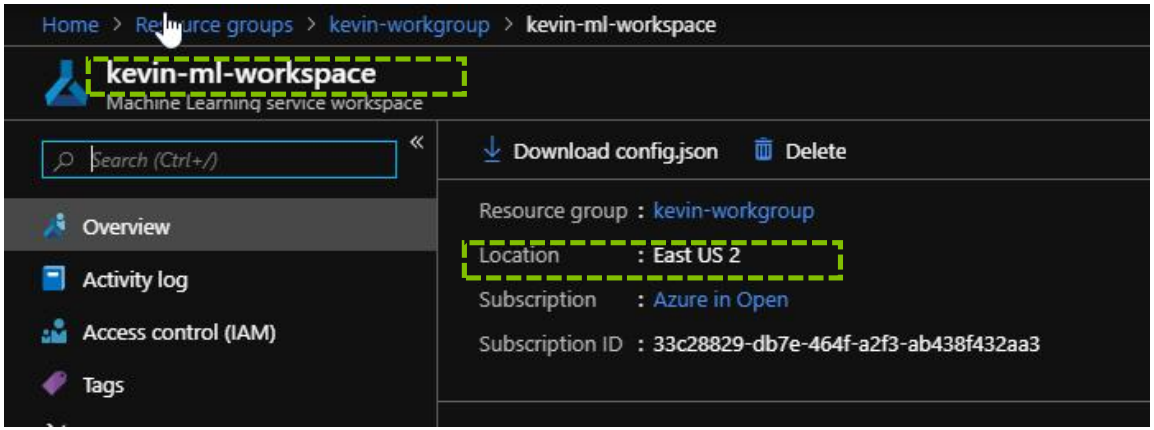
- Login to Azure Notebook -- http://notebooks.azure.com/
- Import sample notebooks into Azure Notebooks
- Click on the Import button.
  - Clear the **Public** checkbox
- Open configuration.ipynb



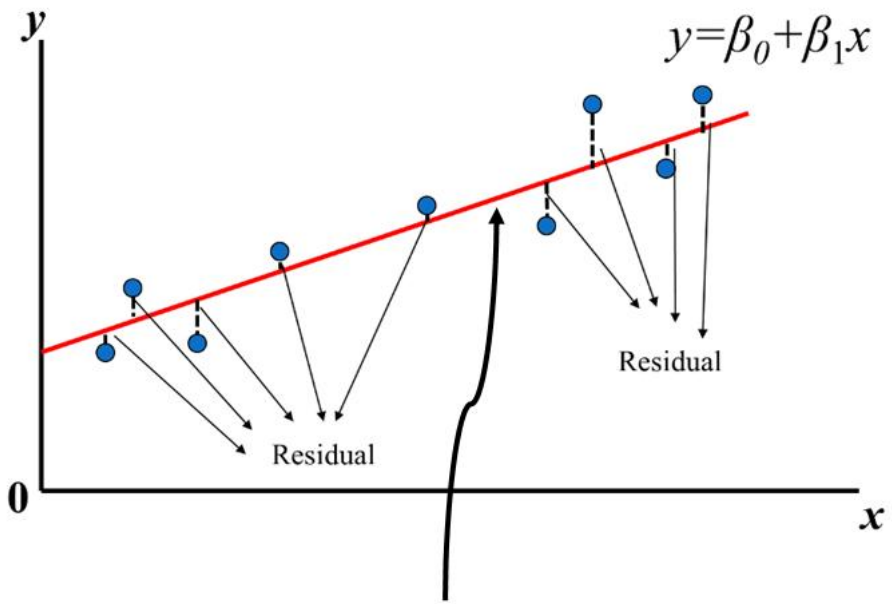
Train an image classification model with Azure Machine Learning

- Open tutorials/img-classification-part1-training.ipynb
  - Set up your development environment
  - Access and examine the data
  - Train a simple logistic regression model locally using the popular scikit-learn machine learning library
  - Train multiple models on a remote cluster
  - Review training results, find and register the best model


Microsoft Azure
Home > Resource groups > kevin-workgroup
Resource groups
WNC Workshop
Add
Edit columns
More
NAME
kevin-workgroup
kevin-workgroup
Resource group
Overview
Activity log
Access control (IAM)
Tags
Events
Settings
Quickstart
Deployments
Policies
Properties
Locks
Export template
Add
Edit columns
Delete resource group
Subscription (change) : Azure in Open
Subscription ID : 33c28829-db7e-464f-a2f3-ab438f432aa3
Tags (change) : Click here to add tags
5 items
Show hidden types
NAME
kevinmlworkspa1201107236
kevinmlworkspa7014182469
kevinmlworkspa7621952264
kevinmlworkspa8248287037
kevin-ml-workspace

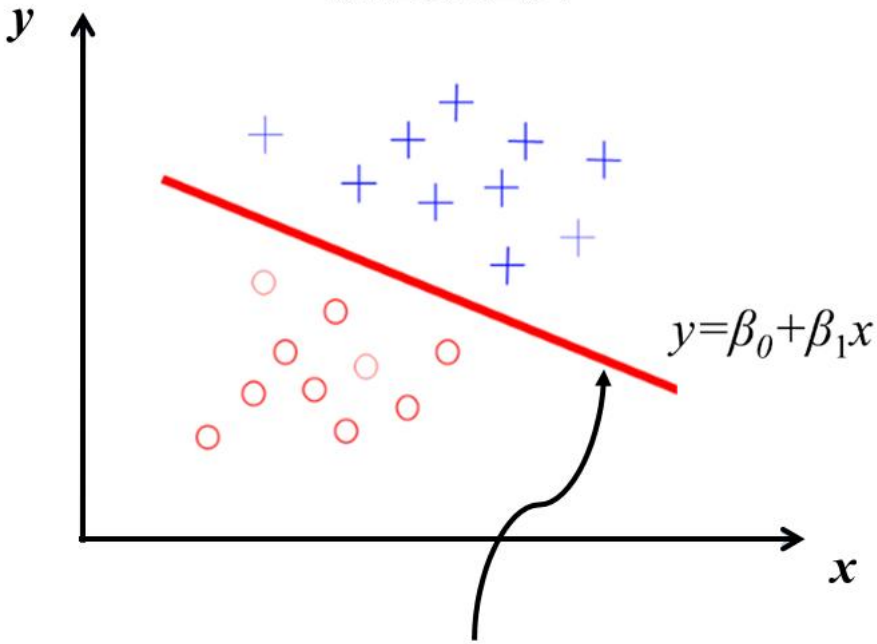

Home > Resource groups > kevin-workgroup > kevin-ml-workspace
kevin-ml-workspace
Machine Learning service workspace
Search (Ctrl+/)
Overview
Activity log
Access control (IAM)
Tags
Download config.json
Delete
Resource group : kevin-workgroup
Location : East US 2
Subscription : Azure in Open
Subscription ID : 33c28829-db7e-464f-a2f3-ab438f432aa3

線性回歸



線性回歸是希望
「找到資料都可以盡量fix的那條紅線」

羅吉斯回歸



羅吉斯回歸希望
「找到那條紅線，讓資料可以區隔開來」

One Hot Encoding

類別資料的處理（有序、無序）

ID	Gender
1	Male
2	Female
3	Not Specified
4	Not Specified
5	Female

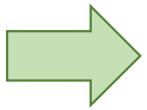
ID	Male	Female	Not Specified
1	1	0	0
2	0	1	0
3	0	0	1
4	0	0	1
5	0	1	0

Categorical Data(類別資料處理)

In [13]:
```
df2 = pd.DataFrame(
    [['green', 'M', 10.1, 1],
    ['red', 'L', 13.5, 2],
    ['blue', 'XL', 15.3, 1]]
)
df2.columns = ['color', 'size', 'price', 'classlabel']
df2
```

Out[13]:

	color	size	price	classlabel
0	green	M	10.1	1
1	red	L	13.5	2
2	blue	XL	15.3	1



In [14]:
```
size_mapping = {
    'XL':3,
    'L':2,
    'M':1
}
df2['size'] = df2['size'].map(size_mapping)
df2
```

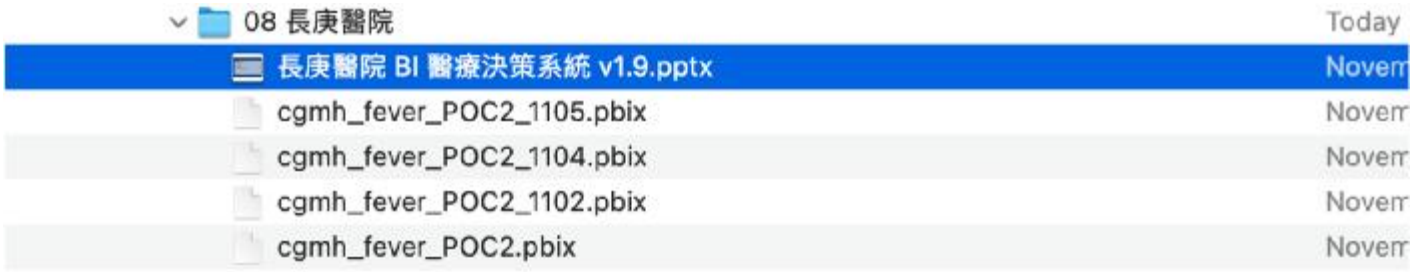
Out[14]:

	color	size	price	classlabel
0	green	1	10.1	1
1	red	2	13.5	2
2	blue	3	15.3	1

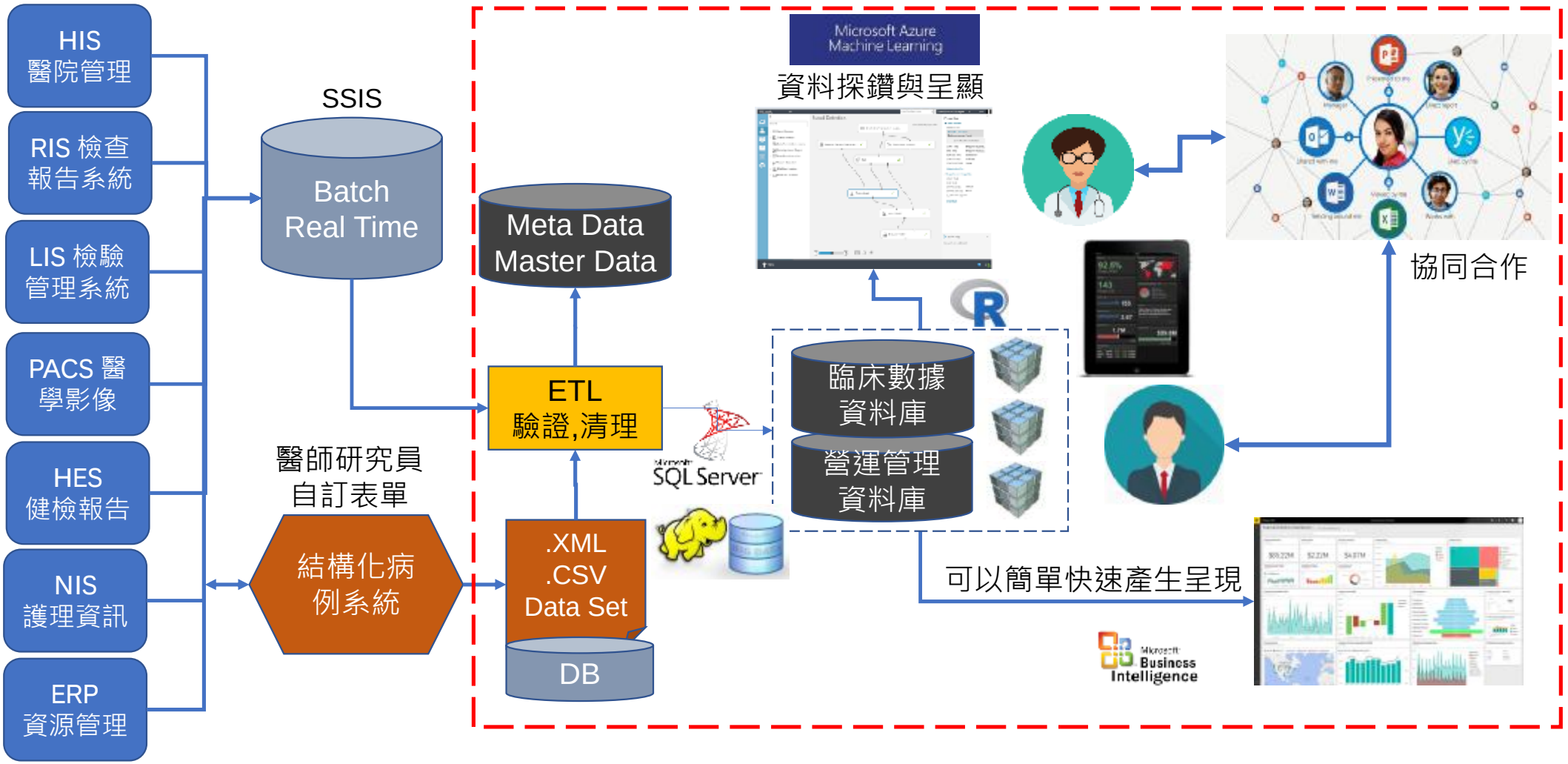
Azure Cloud/Office 365技術經驗

長庚醫院ＢＩ醫療決策系統PoC

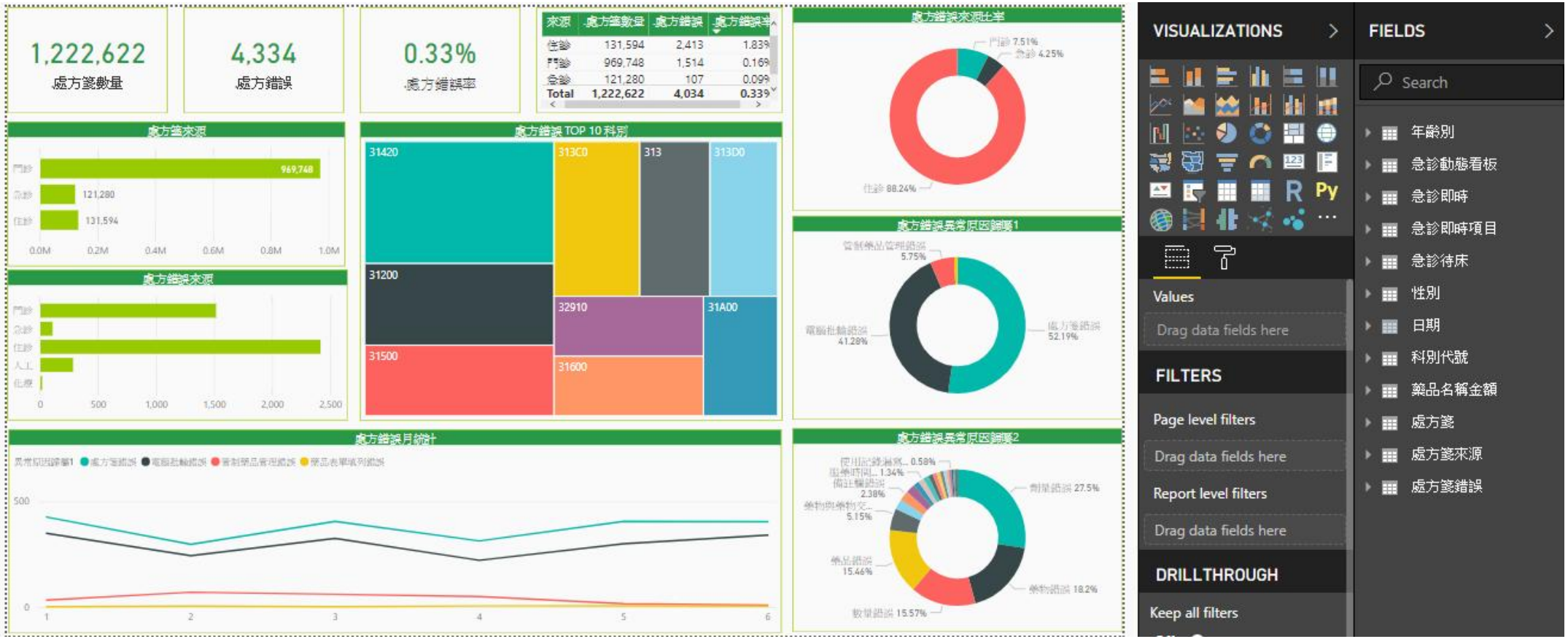
PowerBI 實作Dashbaord的PoC，可於PowerBI雲端以及Desktop執行



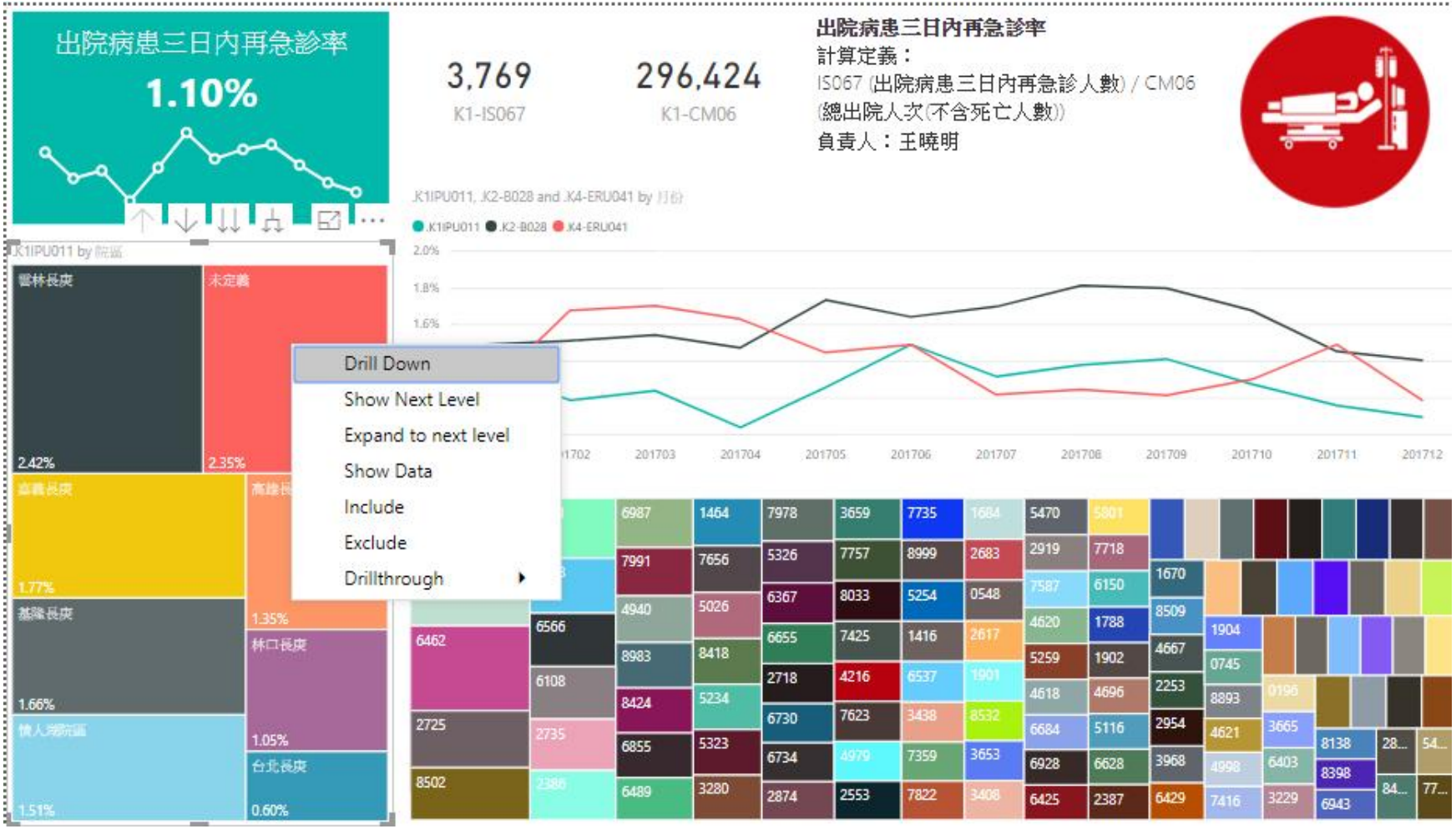
BI 醫療決策分析系統 平台架構



Power BI實作 – 簡易拖拉設計



Power BI 實作- 上鑽下探



Azure Cloud/Office 365技術經驗

Digital Link運用O365 Teams的產品設計

- **目標與效益**

  品質監控。

- **技術架構**

  - **Data Streaming技術，採用微軟Azure雲端個多種服務，（詳如下圖）**

VS.net + SSIS + SSRS + SSAS: 設計ETL以及靜態報表、多維度報表分析。

- **專案角色：開發AngularJS**

Project Goal and Background

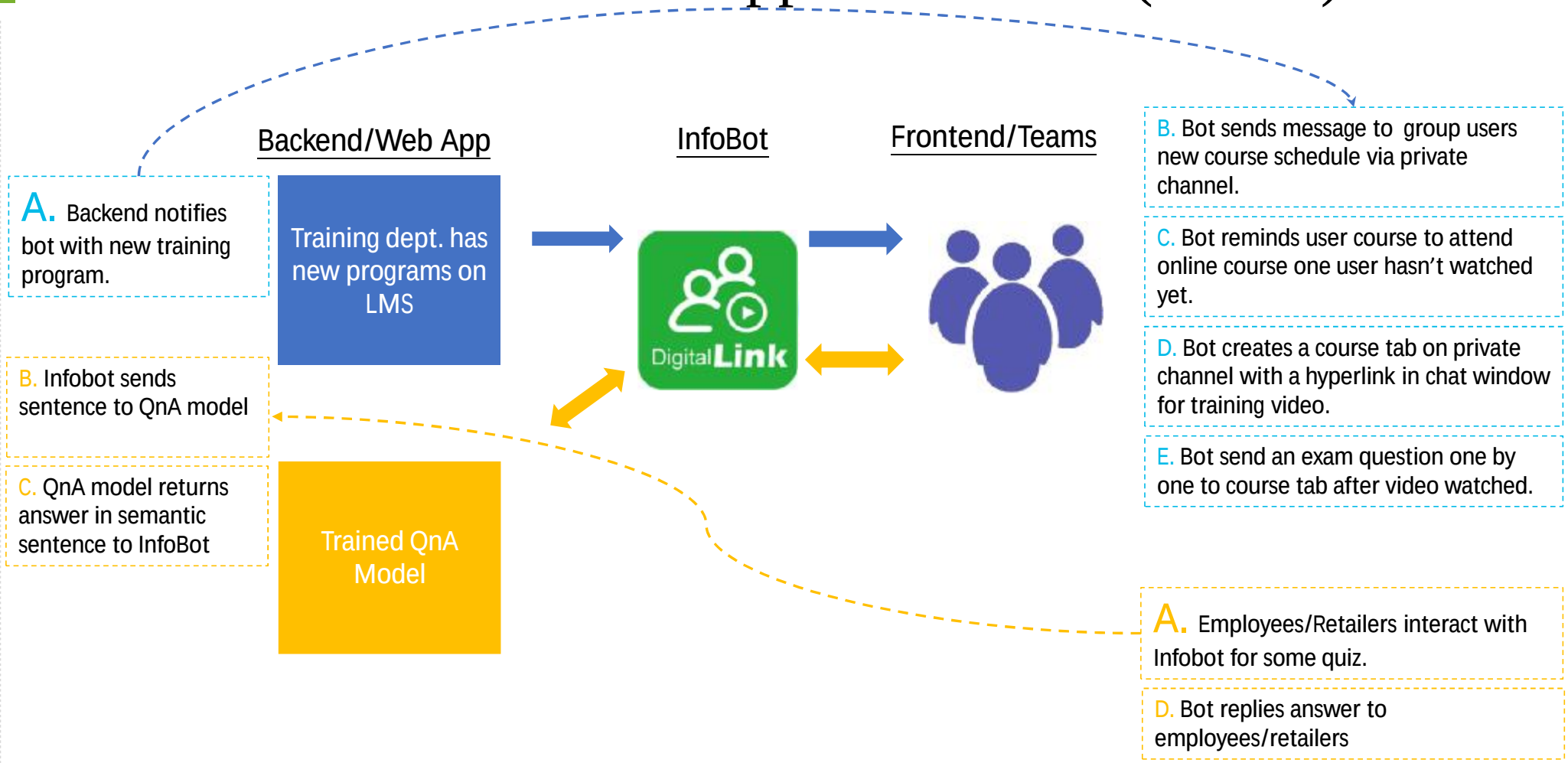
Background	Goal	TA

• Announcement
  • Headquarter
  • Time limited offer
• Online training video
• Chat based QnA with NLP enabled
• Micro exam after course study
	• 1ST version: Hotspot features for IP co-sell
	• 2nd version: Full features for TA(Next project)
		• 美妝櫃與總部
		• 零售業與總部
		• 製造業DL培訓

Advantages

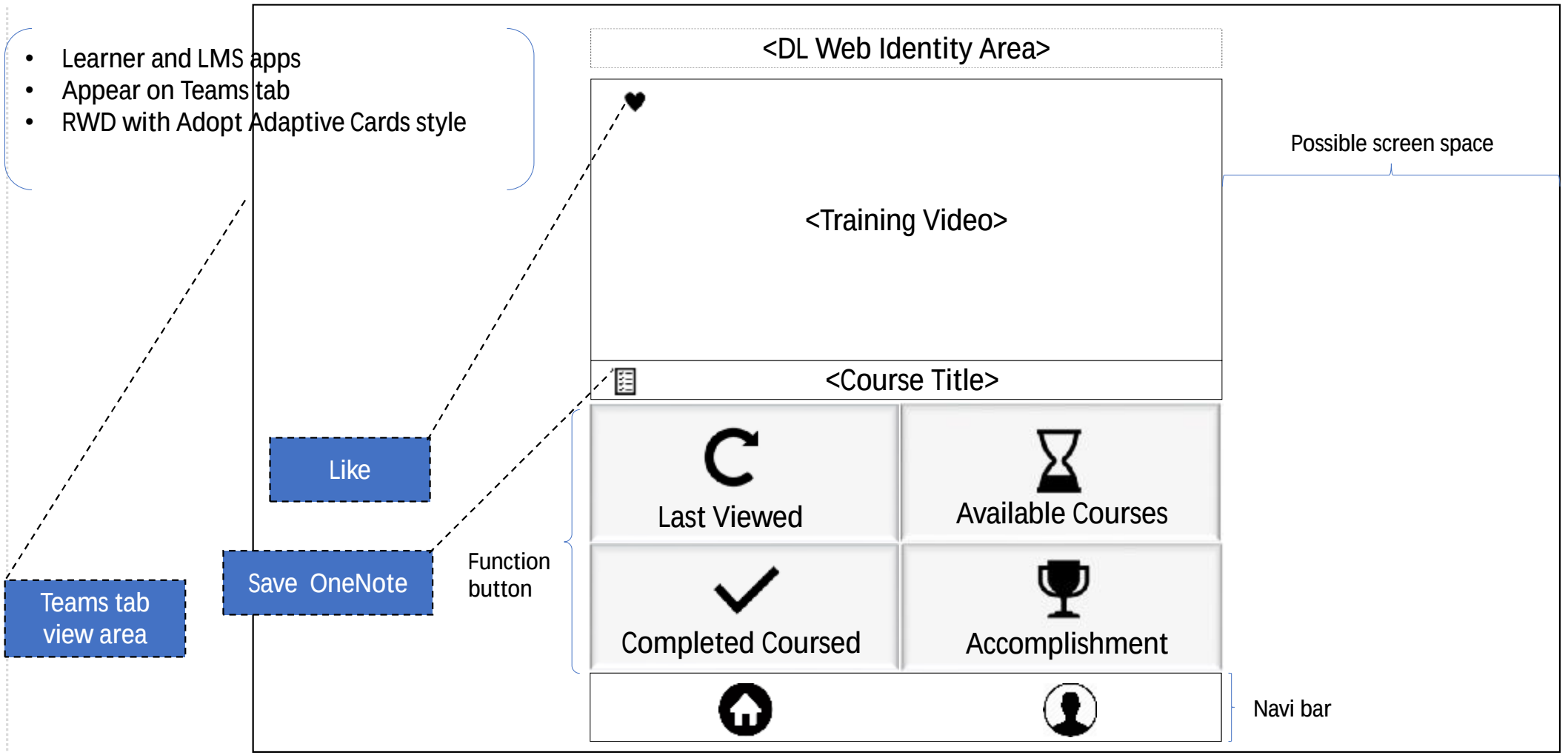
- Common “Announcement features” send notifications to chat window on Teams, schedules to individual’s calendar from headquarter and any approval workflow systems, especially ERP system.
- Chat based QnA with NLP enabled is appropriate for any unmanned CRM, especially for online store, online technical service.

DL Platform and Teams App Interactions(details)



Wireframe - Teams App – Online Training

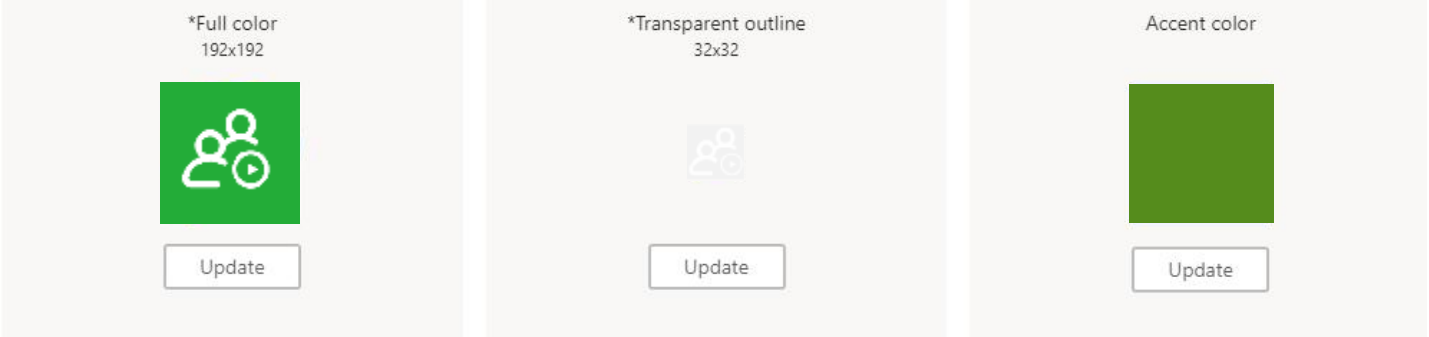
- Learner and LMS apps
- Appear on Teams tab
- RWD with Adopt Adaptive Cards style

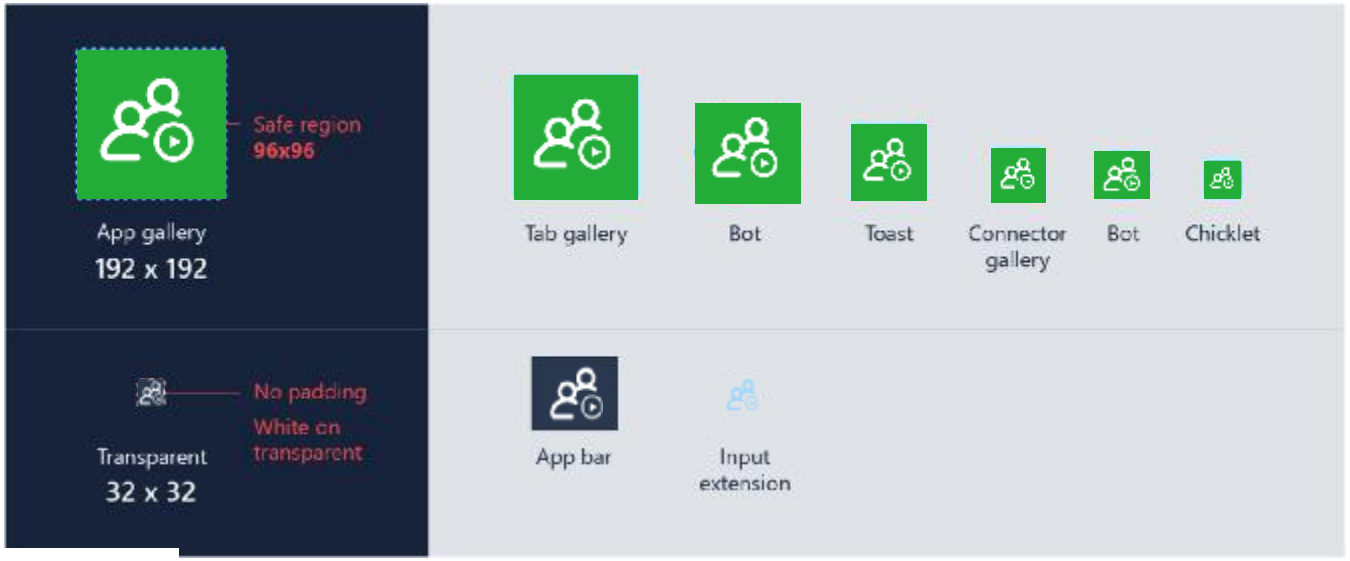


Teams bot Identity (not final ver.)

Digitalink.zip

Bot Manifest





Apply area
Read instruction for more @
https://docs.microsoft.com/en-us/microsoftteams/platform/concepts/apps/apps-package

Upload for UT test
https://docs.microsoft.com/en-us/microsoftteams/platform/concepts/apps/apps-upload

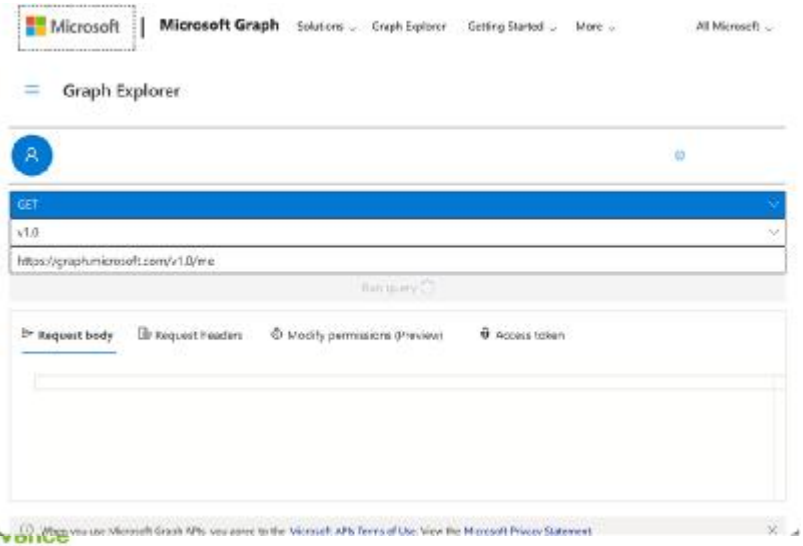
Office 365 Teams開發經驗

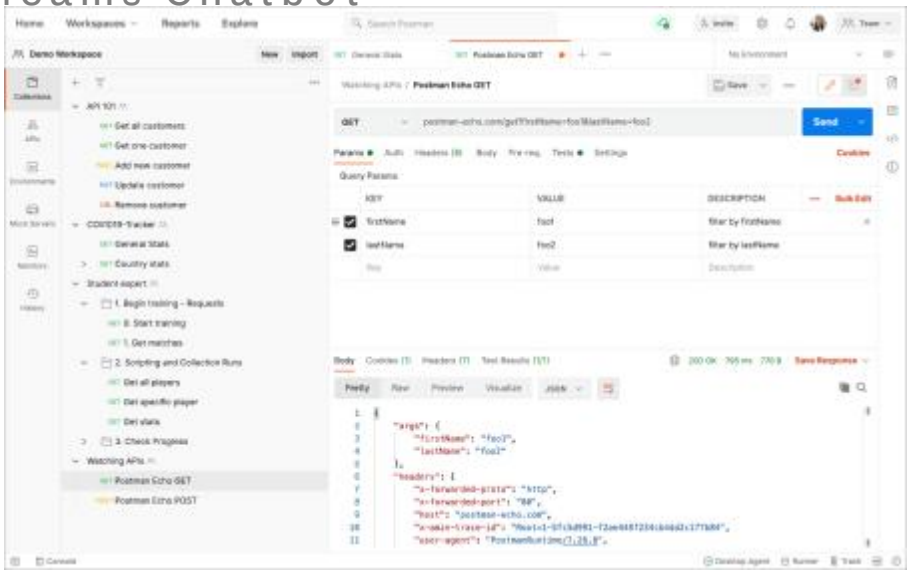
開發工具：Visual Studio Community

使用技術：Azure AD + OAuth, .Net Core Asp.net, NodeJS,

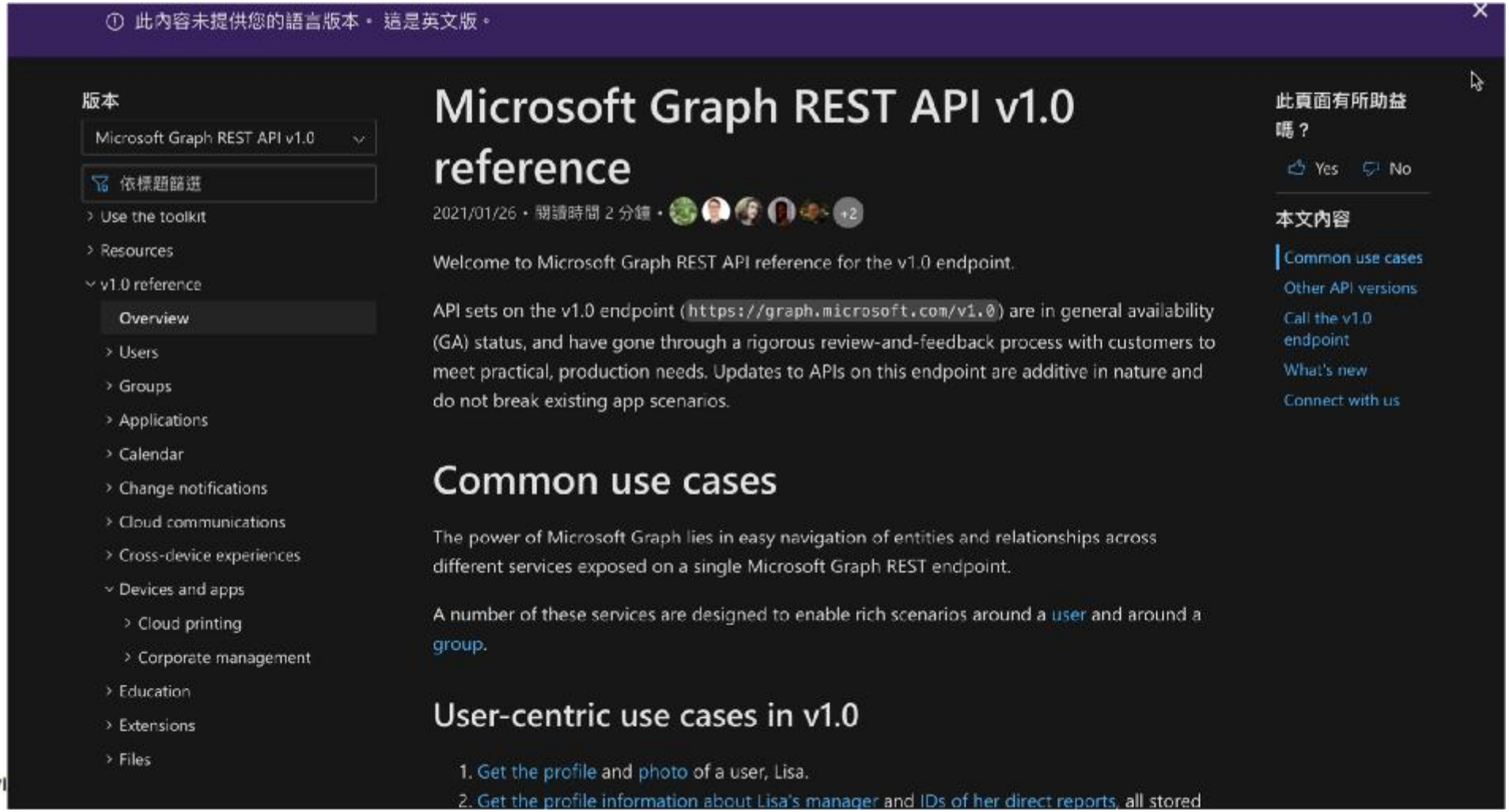
利用Microsfot Graph Explor, Postman, Fiddler發送Http Get, Post測試Graph API功能

技術框架：.Net Core, NodeJS, MS Bot Framework, MS Office 365 Graph API, Azure QnAMaker(LUIS服務）開發Teams Chatbot





Microsoft Graph REST API



AWS技術經驗

equuscs.com, cliqstudios.com

- **目標與效益**

  維運B2B, B2C網站、自動化佈版, QA測試自動化

- **技術架構**
  - 採用AWS VM運行equuscs.com 於Windows 2k6, cliqstudios.com 於ubuntu
  - **建立自動化佈板機制，運用Jenkins串連JIRA, Github, 以python, ansible, powershell script建立自動化CICD佈版。Python與Docker建立自動化測試**

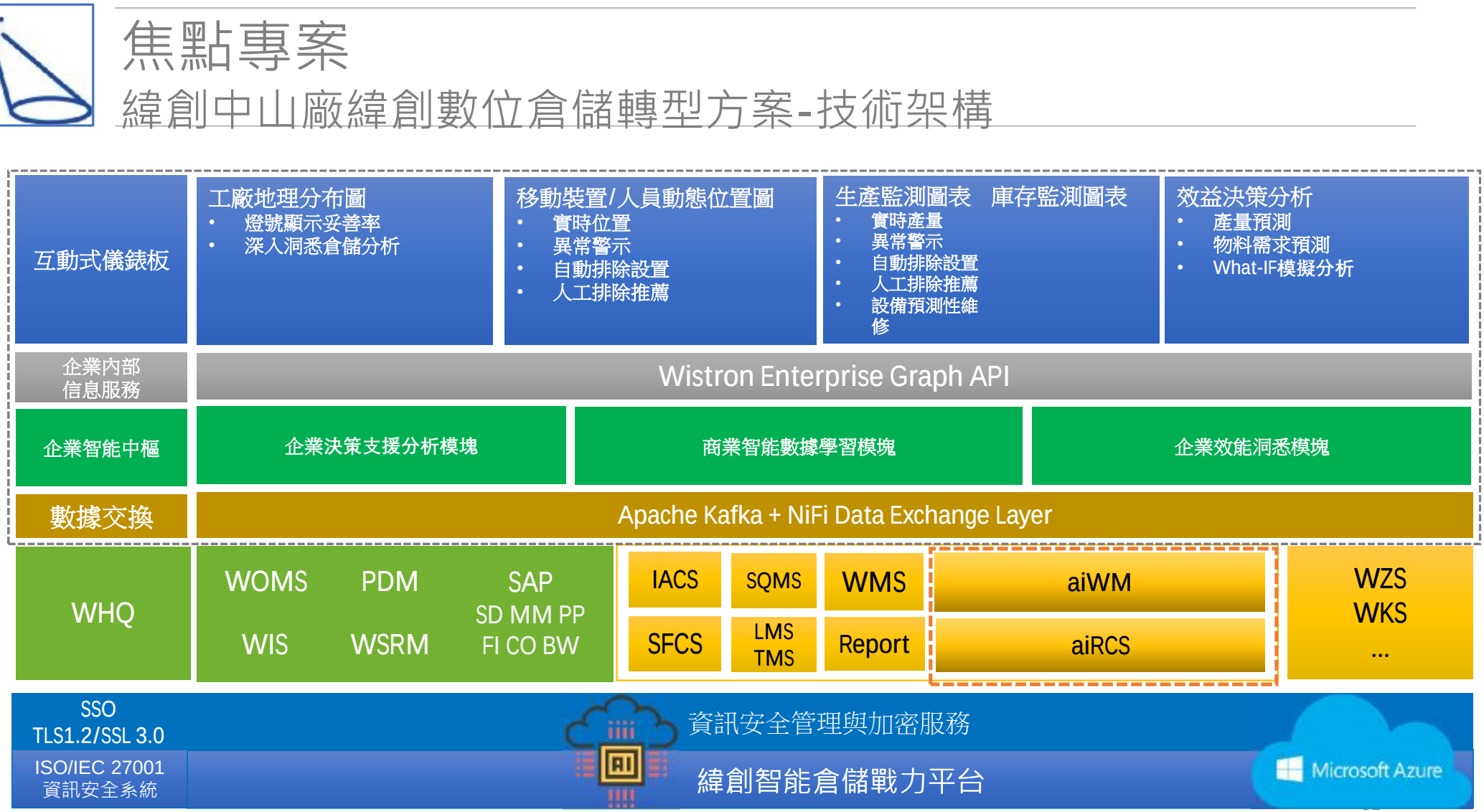
焦點專案

緯創中山廠緯創數位倉儲轉型方案

- **目標：倉儲數位轉型能力**
  - 提升緯創全球倉儲智能
  - 倉儲營運與管理自主精進
  - 緯創全球倉儲即時營運信息
- 效益：
  - 數據量化管理、智能決策支援
  - 流程自動化
  - 量化標準作業程序

焦點專案

緯創中山廠緯創數位倉儲轉型方案-技術架構



焦點專案

華邦電子半導體生產良率分析

- **目標與效益**

  處理與可查詢半年內的巨量生產紀錄(每秒計)

- **技術架構**

  VS.net + SSIS + SSRS: 即時收取巨量生產紀錄由每秒到平均每五分鐘，

  數位儀表板: 開發良率分析數位儀表板，提供生產月份區間以及觀察指標紀錄查詢將查詢結果產生Excle分析報表檔案

焦點專案

消防署防救災緊急應變 EMIC 2.0 系統

- **目標與效益**

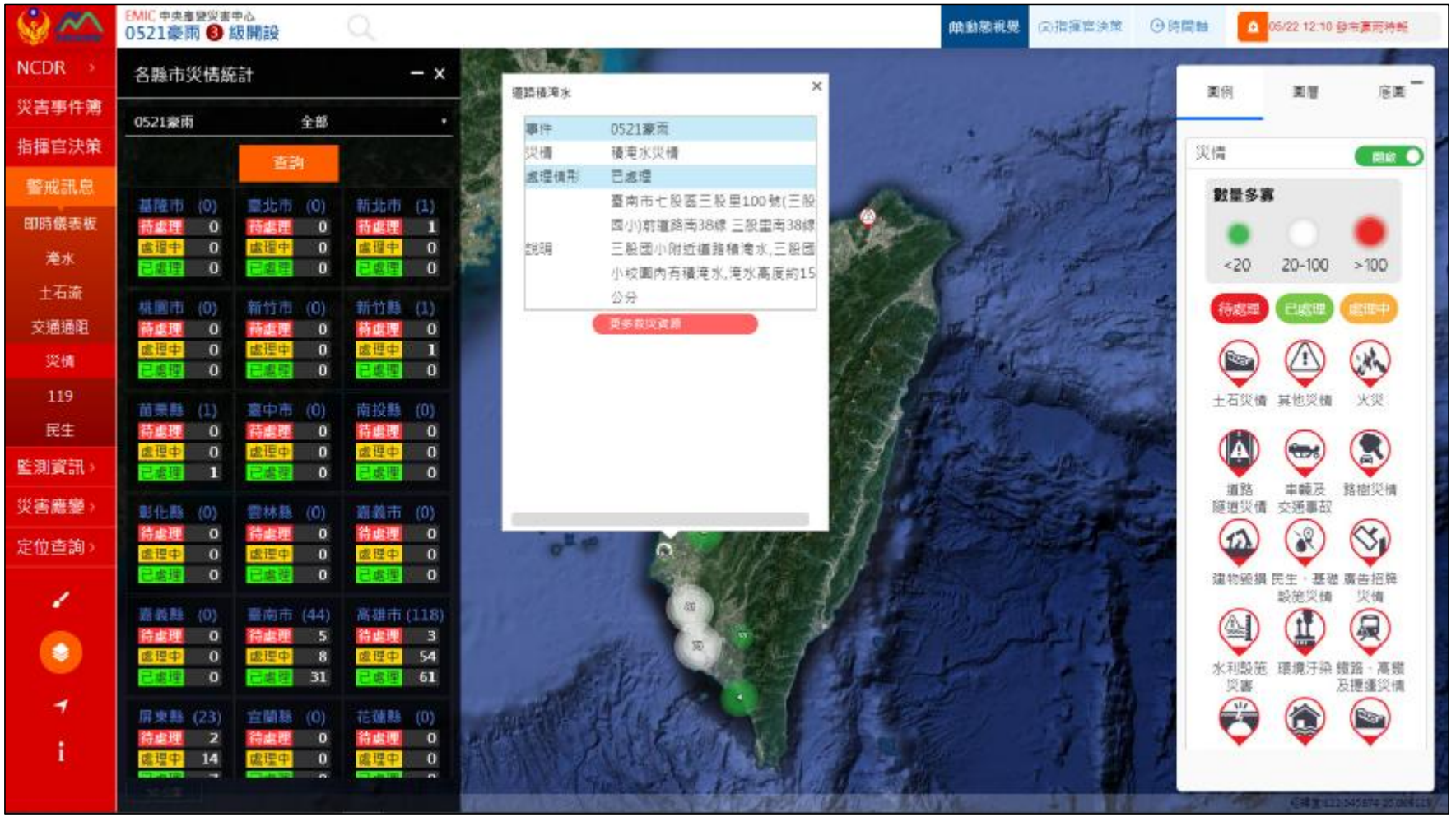
  提高防救災決策支援

- **Technical Architecture**

  VS.net + SSIS + SSAS + SSRS:設計數位儀表板呈現所搜集品質資料狀態、統計防救災設備與器材，一線救災員之效能分析、救災資源庫存狀態

焦點專案

消防署防救災緊急應變 EMIC 2.0 系統



最近專案

於Amazon銷售之防雷擊之智慧排插-

https://www.amazon.com/gp/product/B00TBDXO5K

- 領導智慧排叉之雲平台、App開發從概念、研發與製作
- Manage IoT cloud and devices, iOS app development with JIRA, Mantis
- Java web app on Azure VM with SQL Service
- Integrated Azure IoT Suite and iOS app
- Manage CI/CD DevOps on AWS with Jenkins, Chocolate (Python, Powershell Script, Ansible)
- Analyze marketing segments, sales performance, online after-sales service and RMA in multidimension cubes developed by SSIS, SSRS and SSAS

最近專案

Azure 雲端專案採用Azure 網頁應用Web App、AzureML機器學習服務、MLOps機器學習部署自動化以及Azure LUIS 客服機器人應用

- 高雄市新衛中心智慧心理健康雲補助案
- 緯創中山廠智慧倉儲與AGV搬運機器人管理
- 南亞塑膠樹脂廠之鍋爐節能AI研究案
- 南亞塑膠樹脂廠之ＡＩ生產排程系統
- **Digitalinks 應用微軟O365 Teams開發移動教育訓練App**

  設計微學習、微訓練、為溝通平台

(Next…)

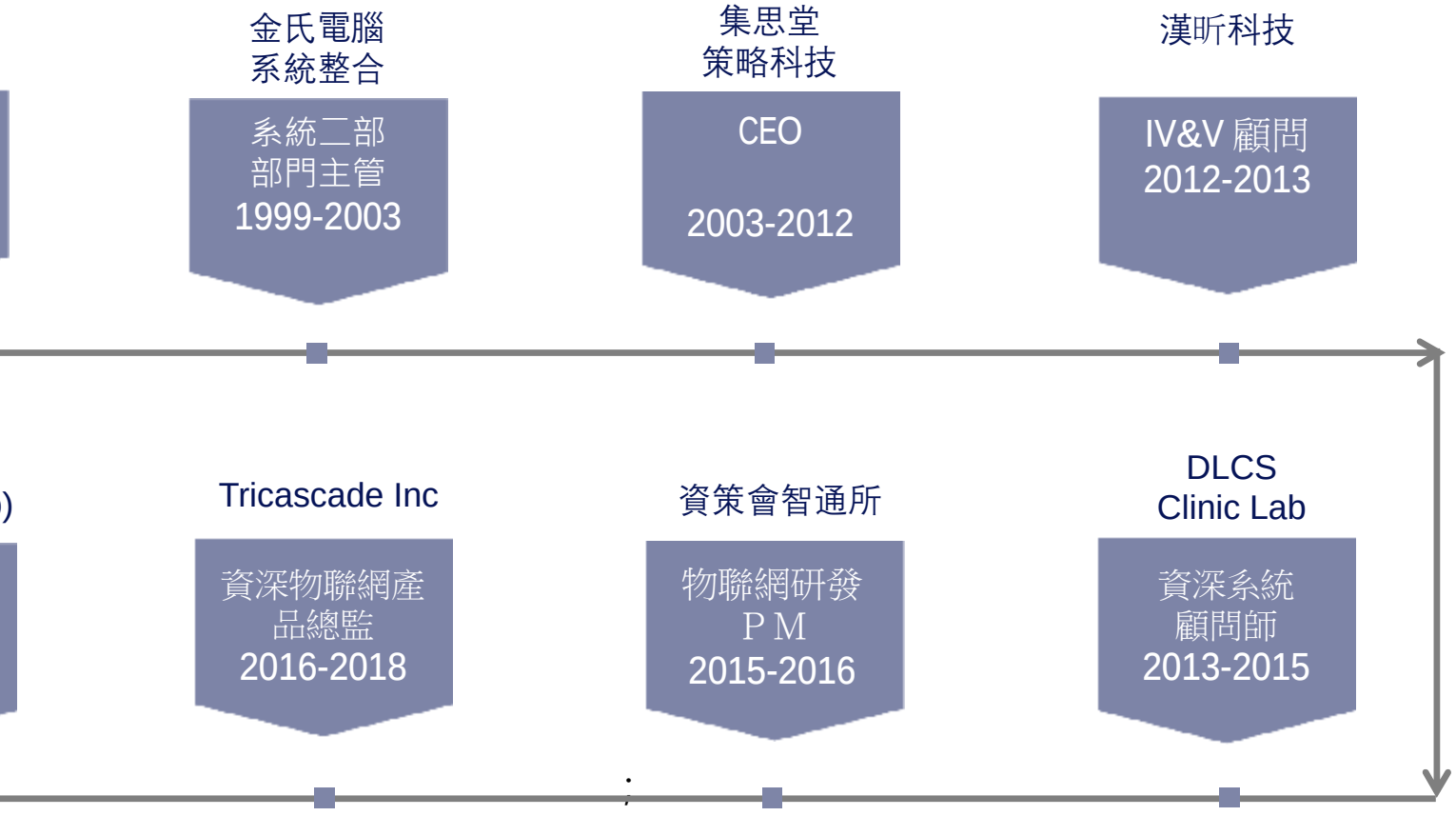
管理能力

- PMP方法論專案管理
- 具敏捷式開發經驗如：Scrum、Pair Programming、eXtreme Programming
- 具有跨功能團隊協作與溝通能力
- 依據企業商業導向，制定IT策略支持決策與
- 依據SMART原則制定目標

確保交付有品質的專案管理方式

<table>
<tr><td>管理方法論</td><td colspan="2">評估</td><td colspan="2">設計</td><td colspan="2">建構</td><td colspan="2">實作</td><td colspan="2">營運</td></tr>
<tr><td>軟工生命週期</td><td colspan="2">規劃與系統需求分析</td><td colspan="2">系統設計</td><td colspan="2">開發</td><td colspan="2">測試與整合</td><td colspan="2">運營維護</td></tr>
<tr><td rowspan="9"></td><td>輸入</td><td>輸出</td><td>輸入</td><td>輸出</td><td>輸入</td><td>輸出</td><td>輸入</td><td>輸出</td><td>輸入</td><td>輸出</td></tr>
<tr><td>初步用戶需求</td><td>專案報價</td><td></td><td></td><td></td><td rowspan="8">程式開發與單元測試</td><td></td><td></td><td rowspan="8">具品質的系統</td><td></td></tr>
<tr><td></td><td>專案管理計畫</td><td>專案管理計畫</td><td>發展測試計畫</td><td>測試計畫</td><td>測試計畫</td><td>SIT</td><td>Bug回報</td></tr>
<tr><td></td><td>定義範疇</td><td rowspan="5">Scope, WBS, schedule & project members</td><td>Sys. infra.</td><td>Sys. infra.</td><td></td><td>UAT</td><td rowspan="5"></td></tr>
<tr><td></td><td>發展WBS</td><td>Sys. modules</td><td>Sys. modules</td><td></td><td rowspan="4"></td></tr>
<tr><td></td><td>安排時程</td><td>Sys. objects</td><td>Sys. objects</td><td></td></tr>
<tr><td></td><td>配置人力：SA, SD, DBA, PG & QA</td><td>Sys. APIs</td><td>Sys. APIs</td><td></td></tr>
<tr><td></td><td>準備專案管理工具 (PM, 版控...etc)</td><td></td><td></td><td></td><td></td></tr>
<tr><td></td><td>需求規格書 (SRS)</td><td>SRS</td><td>SDS</td><td>SDS</td><td></td><td>交付具品質的系統</td><td>發現新需求</td></tr>
</table>

工作經驗//25年以上


方向國際設計
Operation Director 1995-1999
金氏電腦系統整合
系統二部部門主管 1999-2003
集思堂策略科技
CEO 2003-2012
漢昕科技
IV&V 顧問 2012-2013
緯謙科技 (Wistron Group)
前瞻技術部主管 2018-2019
Tricascade Inc
資深物聯網產品總監 2016-2018
資策會智通所
物聯網研發 P M 2015-2016
DLCS Clinic Lab
資深系統顧問師 2013-2015

專業技能

- 專案管理工具：Jira, Redmine 2000, Mantis, Azure Devops Boards, Excel, Project 365, Trello
- 商業智慧方案： Power BI desktop, VS.net with SSIS, SSRS and SSAS
- 機器學習： Azure ML service, Anaconda(JupyterLab, Spyder)
- 程式語言與框架：
  - .Net framework 4.0+ C#, Python, Javascript (Angular, Jquery, Node)
- **RDBMS/NoSql**：MS/Azure SQL 2019, Oracle 9i, MongoDB
- **Data Streaming**：Apache Kafka + NiFi
- 資料庫設計工具：ER-Win, MS Visio, PowerDesigner
- 物件導向工具：Enterprise Architect, MS VISIO

教育與認證

- 大同大學資訊經營所碩士(**榜首**) 主修商業智慧 2010 / 08 ~ 2015 / 01畢
- 論文發表：應用擇優策略粒子群結合區域搜索演算法於資料分群問題 2014/12
- PMP 教育訓練 2007/11
- 商業智慧規劃師 2011/03
- AI人工智慧學校經理人班第五期結業 2019/11

更多專案經歷

- 大同大學資訊經營所碩士(**榜首**) 主修商業智慧 2010 / 08 ~ 2015 / 01畢
- 論文發表：應用擇優策略粒子群結合區域搜索演算法於資料分群問題 2014/12
- PMP 教育訓練 2007/11
- 商業智慧規劃師 2011/03
- AI人工智慧學校經理人班第五期結業 2019/11