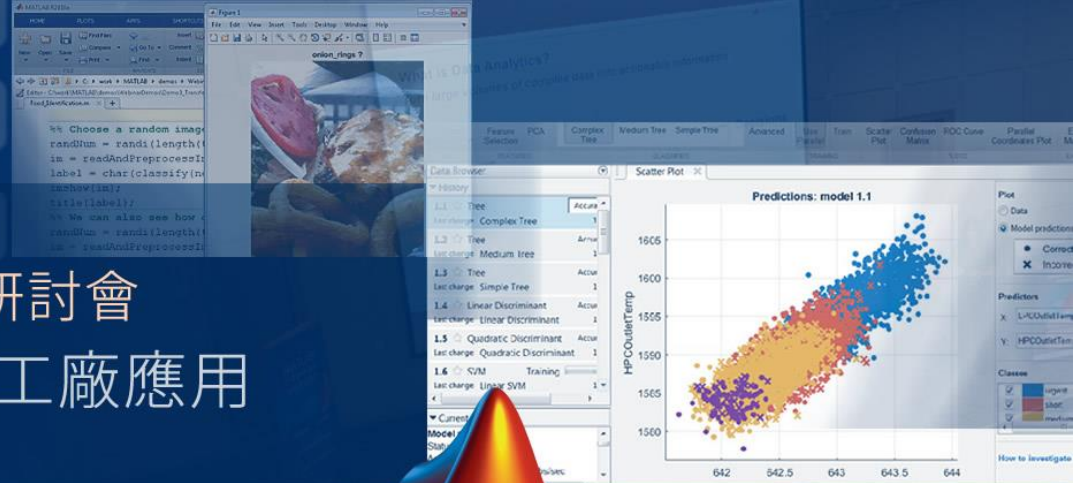


2021 MATLAB校園巡迴研討會

人工智慧、程式技巧、智慧工廠應用



從前端資料擷取到後端文本/AI分析： MATLAB與AIoT開發整合流程

Internet of Things

應用層

雲端主機接受網路層所傳遞資訊，藉由大數據分析與機器學習等演算法處理資料生成有用解決辦法。



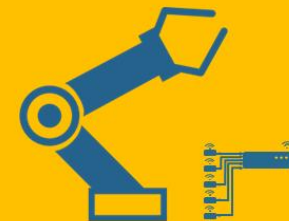
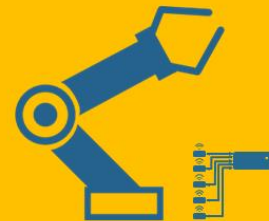
網路層

資訊透過有線或 3G、wi-fi 等無線通訊科技傳遞資訊。

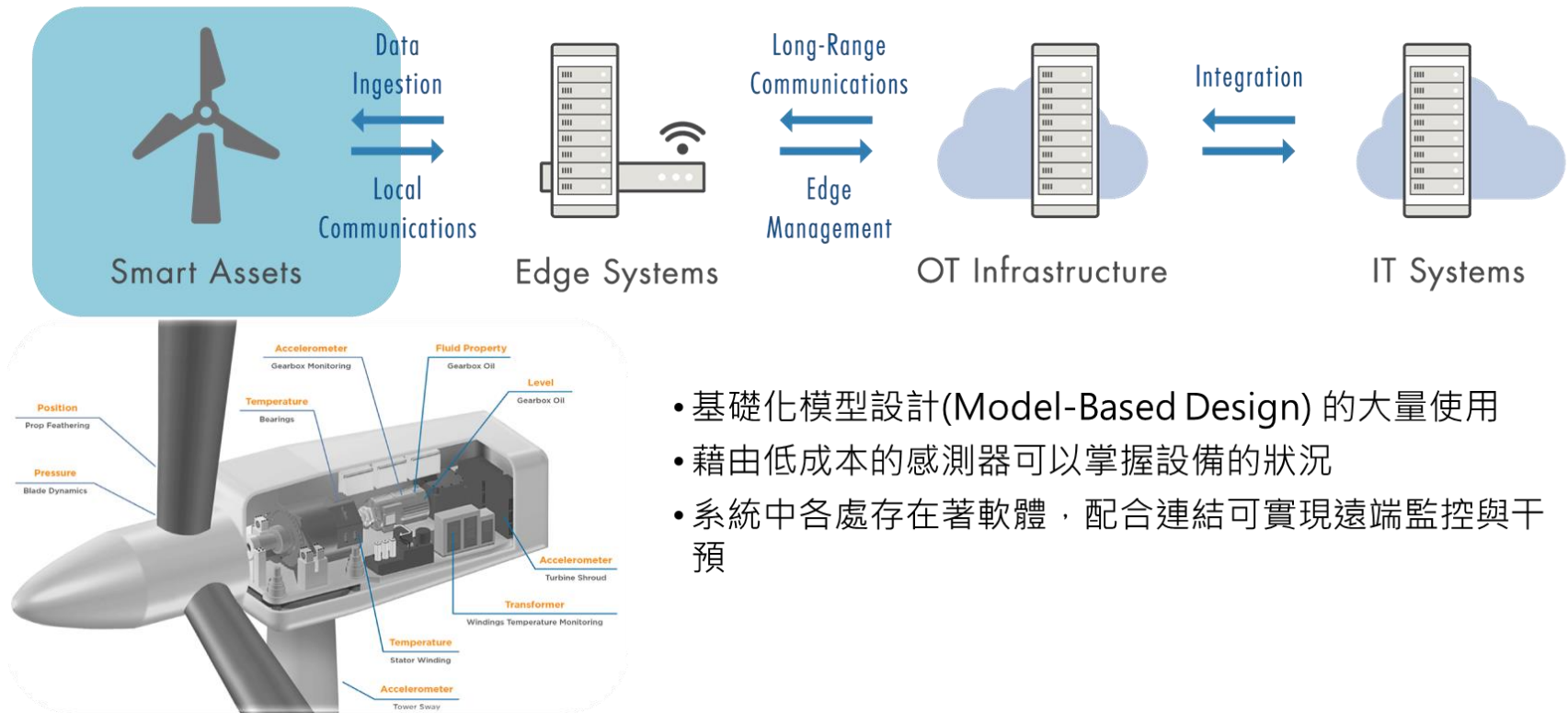


感測層

透過壓力、扭力、濕度、加速規、鏡頭等各式感測計收集機台或是流程數據。

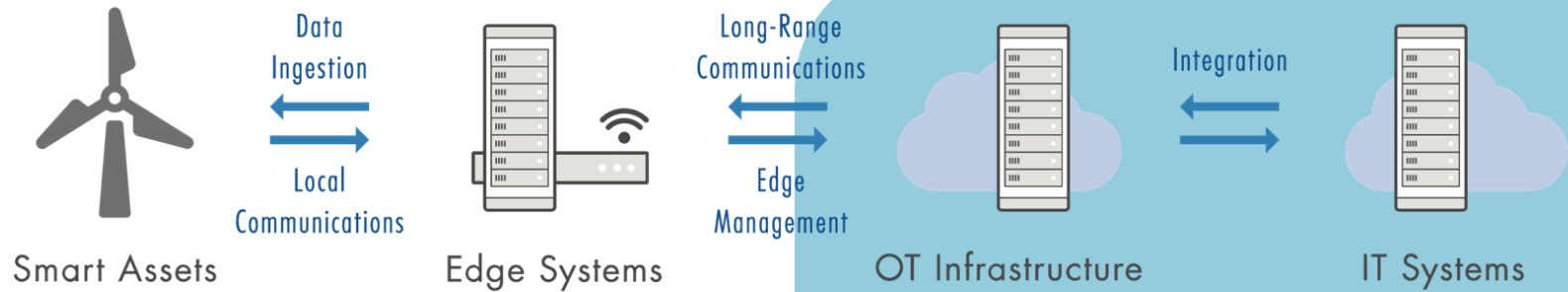


Industrial IoT Systems (IIoT)



- 基礎化模型設計(Model-Based Design) 的大量使用
- 藉由低成本的感測器可以掌握設備的狀況
- 系統中各處存在著軟體，配合連結可實現遠端監控與干預

Industrial IoT Systems



Enterprise Integration Focus

資料串流, 第三方整合,
雲端整合/執行, 應用產品佈署



Data Aggregation

- Preprocessing, Synchronization
- Database Accesses



Edge Devices

- Signal Processing, Data Compression
- Embedded Systems (CPU, FPGA)

“A” IoT



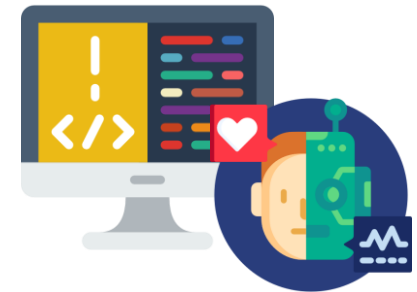
Data Aggregation

- Preprocessing, Synchronization
- Database Accesses



Edge Devices

- Signal Processing, Data Compression
- Embedded Systems (CPU, FPGA)



Data Analytics

- Trend Analysis, Anomaly Detection
- Data Visualization, Reporting

數位黃金



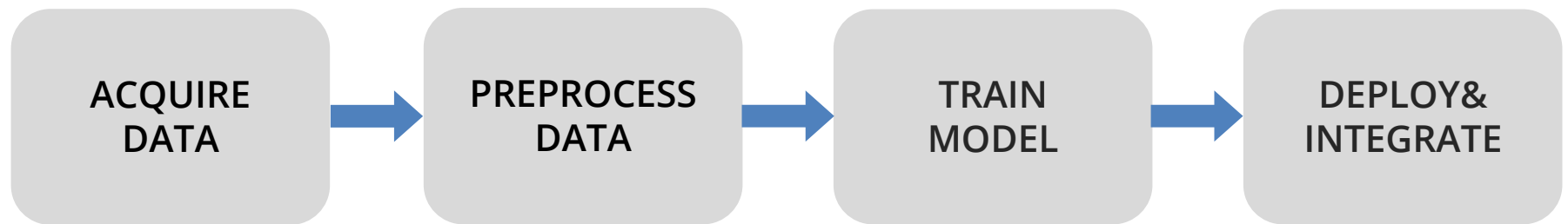
數位黃金

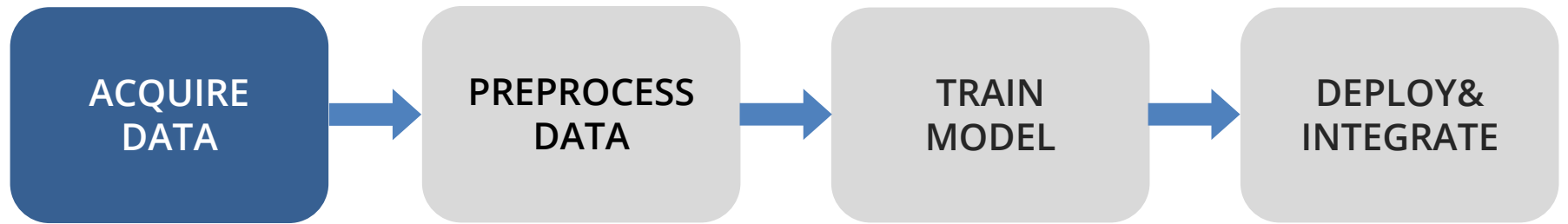


資料何處來 此物只應天上有?



Process workflow





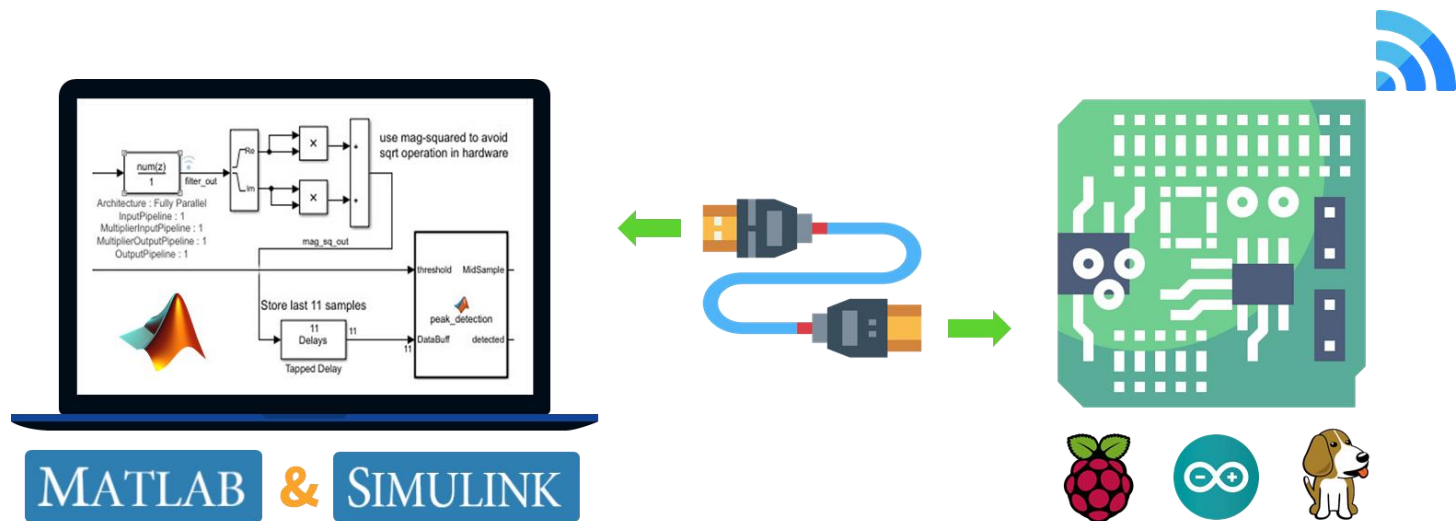
資料收集

當你在 Messenger 群組貼了超爆笑的梗圖時



Source: <https://hornydragon.blogspot.com/2018/04/974.html>

Edge Device



可以將**MATLAB & Simulink**中所開發的演算法部屬到終端硬體中，
用於連結感測器並收集資料

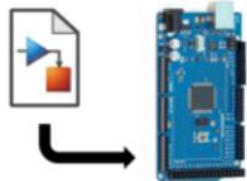
Hardware Library



MATLAB Support Package for Arduino Hardware

by MathWorks MATLAB Hardware Team **STAFF**

Acquire inputs and send outputs on Arduino boards



Simulink Support Package for Arduino Hardware

by MathWorks Simulink Team **STAFF**

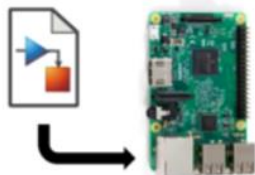
Run models on Arduino boards.



MATLAB Support Package for Raspberry Pi Hardware

by MathWorks MATLAB Hardware Team **STAFF**

Acquire sensor and image data from your Raspberry Pi.



Simulink Support Package for Raspberry Pi Hardware

by MathWorks Simulink Team **STAFF**

Run models on Raspberry Pi.

Arduino + AMT1001

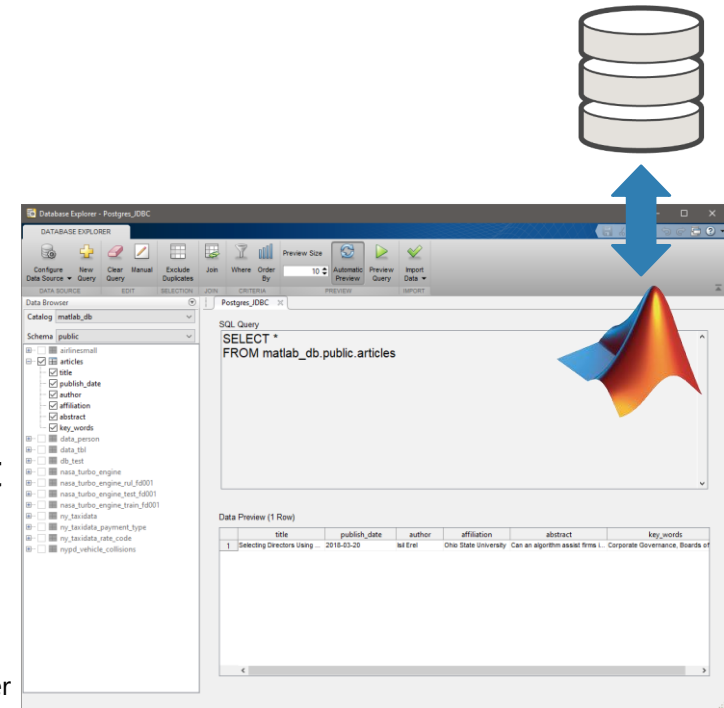
The screenshot displays the MATLAB R2021a environment. The main window shows a script named 'AMT1001_Sensor.m' with the following code:

```
1 %% Arduino with AMT1001 sensor
2 %% Connct to Arduino
3 % Install arduino support packages from Add-on
4 % https://www.mathworks.com/matlabcentral/fileexchange/47522-matlab-support-packag
5
6 clear -global ArduinoObj
7 global ArduinoObj
8 ArduinoObj = arduino('COM4');
9 %% Collect & Convert Data
10 % Raw voltage data from AMT1001 sensor
```

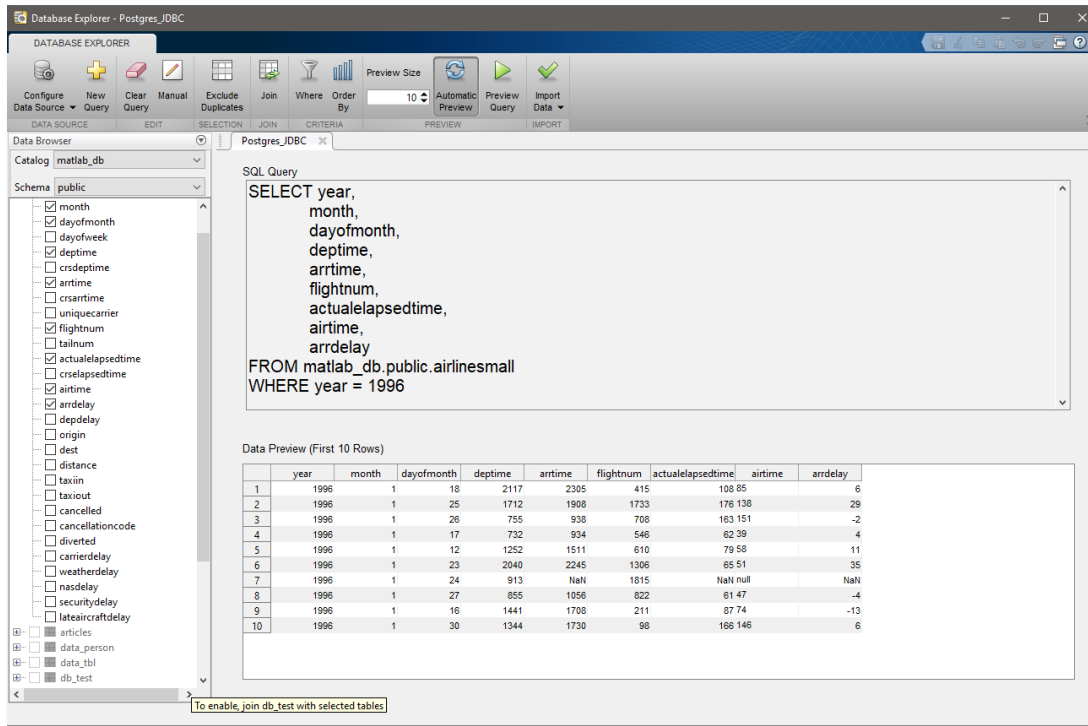
The Command Window shows the prompt `fx >>`. The Workspace and Command History panes are empty. The status bar at the bottom indicates the file is 'AMT1001_Sensor.m (Script)' in UTF-8 encoding, with the cursor at line 6, column 25.

Database

- Connection with relational DB (JDBC / ODBC compliant)
 - Simple command line without SQL
 - GUI access (Database Explorer app)
- Connection with non-relational DB
 - Neo4j (Graph DB)
 - MongoDB (Document Oriented DB)
 - Apache Cassandra (column-oriented DB)
- Reading and writing to SQLite files that do not require DB
- Supports datastore for big data
- SQL query partitioning, parallel access *
※ Required Parallel Computing Toolbox and MATLAB Distributed Computing Server

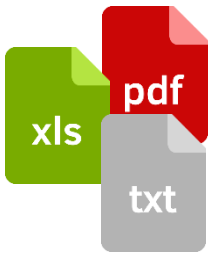


Database Explorer



可以在不了解SQL的情況下從GUI訪問和讀取數據
自動生成用於導入的SQL查詢語句和MATLAB腳本

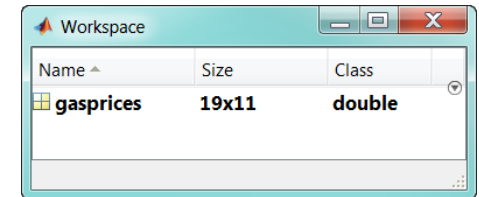
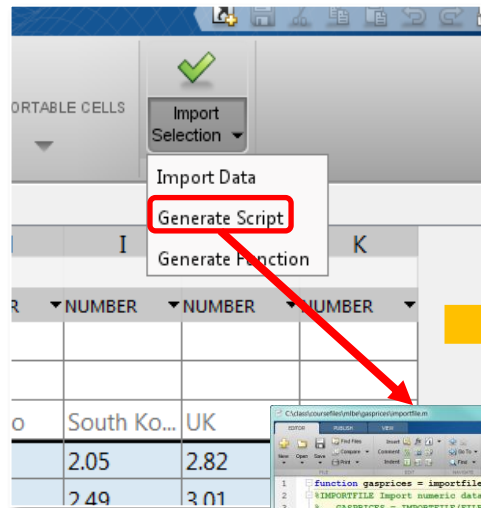
Local Files



Spreadsheet
Delimited text



Audio
Video
Image
Scientific data
Geographic data
...



```
function gasprices = importfile(filename, startRow, endRow)
% IMPORTFILE Import numeric data from a text file as a matrix
% GASPRICES = IMPORTFILE(FILENAME) Reads data from text file
% the default selection.
%
% GASPRICES = IMPORTFILE(FILENAME, STARTROW, ENDRROW) Reads
% STARTROW through ENDRROW of text file FILENAME.
%
% Example:
% gasprices = importfile('gasprices.txt', 4, 22);
% See also TEXTSCAN.
%
% Auto-generated by MATLAB on 2012/07/09 16:36:37
%
%% Initialize variables.
delimiter = '\t';
if nargin<=2
    startRow = 4;
    endRow = inf;
end
%% Read columns of data as strings:
% For more information, see the TEXTSCAN documentation.
formatSpec = '%s%s%s%s%s%s%s%s{\n\}';
```

Import Data



1

Access Data

```
imread  
webread  
parquetread  
...
```

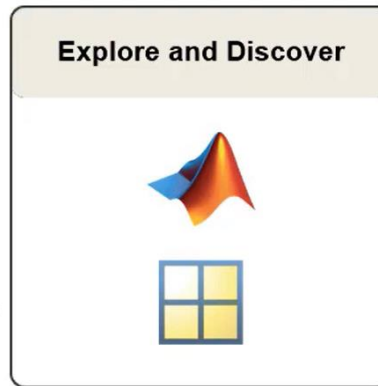
```
imwrite  
webwrite  
parquetwrite  
...
```



***read**
* = FROM
(file, source)

read*
* = INTO
(data type, format)

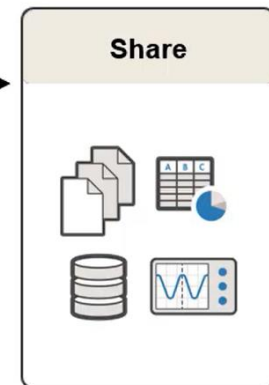
```
readtable  
readmatrix  
readstruct  
...
```



***write**
* = TO
(file, source)

write*
* = OUT OF
(data type, format)

```
writetable  
writematrix  
writestruct  
...
```



Web Crawling



Demo case – Dcard web crawler

Web Data

RESTful Web Service

| | |
|------------|--|
| webread | Read data |
| webwrite | Write data |
| websave | Save data to file |
| weboptions | Specify options such as authentication and timeout |

JSON

| | |
|------------|------------|
| jsondecode | jsonencode |
|------------|------------|

HTTP Messaging


Use the HTTP interface for more complex web communication: ➡➡

```
body = ...
    matlab.net.http.MessageBody(x);

request = ...
    matlab.net.http.RequestMessage ...
    (method,header,body);
```

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|-----------|----------------|-------------------|----------------|----------------|--------|--------------|-------------------|--------------|
| | id | title | excerpt | createdTime | updatedTime | gender | school | department | commentCount |
| 1 | 235969282 | "清水寺附近的親子... | "店名：ひさご親子丼，地... | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "もちもち貝果" | "mochibagel" | 0 |
| 2 | 235968985 | "7-11三起司vs... | "最近這兩個實在有燒到我... | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "國立雲林科技大..." | " | 0 |
| 3 | 235968853 | "不甜不膩的白葡..." | "一直很想做這種清新風的蛋..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "參阿甜" | "abc12300901" | 0 |
| 4 | 235968852 | "不甜不膩的白葡..." | "一直很想做這種清新風的蛋..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "參阿甜" | "abc12300901" | 0 |
| 5 | 235968751 | "#台中美食#手..." | "哈囉！大家好~我們是怪..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "怪獸美食公司" | "monsterfood1..." | 0 |
| 6 | 235968713 | "火鍋料" | "想知道大家在火鍋裡面有什..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "seven" | "sevenseven777" | 1 |
| 7 | 235968616 | "初次手做可愛便..." | "小班校外教學~這是第一..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "Natalia" | "nataliaili" | 0 |
| 8 | 235968605 | "新北淡水 悠哉..." | "淡江學生一定知道這家蛋..." | 2021年5月12日 ... | 2021年5月12日 ... | "F" | "露露肚子餓" | "lu_801" | 0 |
| 9 | 235968592 | "新政府的起司..." | " | 2021年5月12日 ... | 2021年5月12日 ... | "M" | "國立雲林科技大..." | " | 0 |

Demo case – Google comments crawler



所有評論

4.3
6,464 篇評論

撰寫評論

排序

全部 景觀 565 下午茶 275 干貝 228

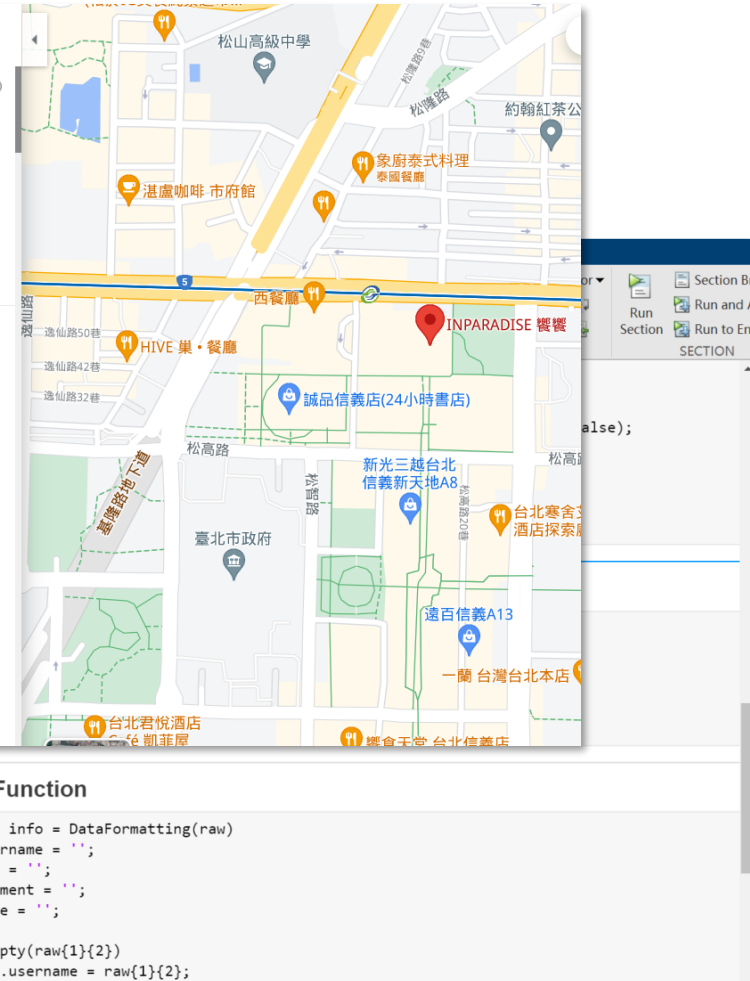
舒芙蕾 203 夜景 184 窗 180

電梯 147 羊排 98

Yvonne
在地嚮導 · 46 則評論

★★★★☆ 3 週前

位在信義微風46樓的吃到飽餐廳
饗饗果然名不虛傳
假日下午茶一個人1290\$/晚餐1890\$
價格雖然很高 但用料實在
將自助餐分為約六區塊



Helper Function

```

61 function info = DataFormatting(raw)
62 info.username = '';
63 info.day = '';
64 info.comment = '';
65 info.rate = '';
66
67 if ~isempty(raw{1}{2})
68     info.username = raw{1}{2};
    
```

ans = 8x4 table

| | username | day | comm... | rate |
|---|------------------|---------|------------|------|
| 1 | 'Yvonne' | '3 週前' | '位在信義微... | 4 |
| 2 | 'Jimmy Wang' | '3 週前' | '人山人海吃... | 2 |
| 3 | 'Drodma Tharl... | '2 個月前' | '饗饗果然名... | 4 |
| 4 | 'Austin Wu' | '1 個月前' | '食材種類非... | 4 |
| 5 | 'Sean Huang' | '1 週前' | '餐點選擇多... | 5 |
| 6 | 'Sunny Farmer' | '1 個月前' | '個人覺得是... | 5 |
| 7 | '曼蒂' | '1 個月前' | '2個月前就要... | 5 |
| 8 | 'KaiDir' | '2 個月前' | '週四晚上六... | 5 |

資料的類型



有點像專有名詞與量詞

屬性資料 (Qualitative Data)

不以數值來表示，以類別來區分資料，又稱為『類別資料』(Categorical Data)。如性別、學歷、公司、國家等等...

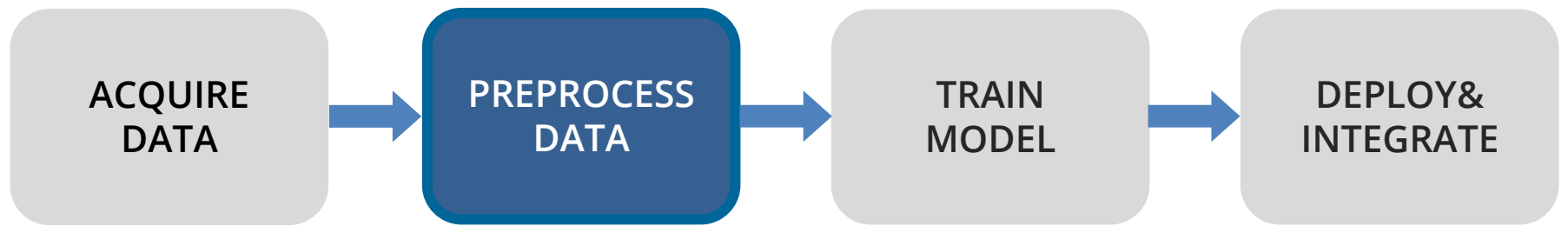
屬量資料 (Quantitative Data)

離散資料 (Discrete Data)

屬於可數的屬量資料，
例如：人數，機器台數等

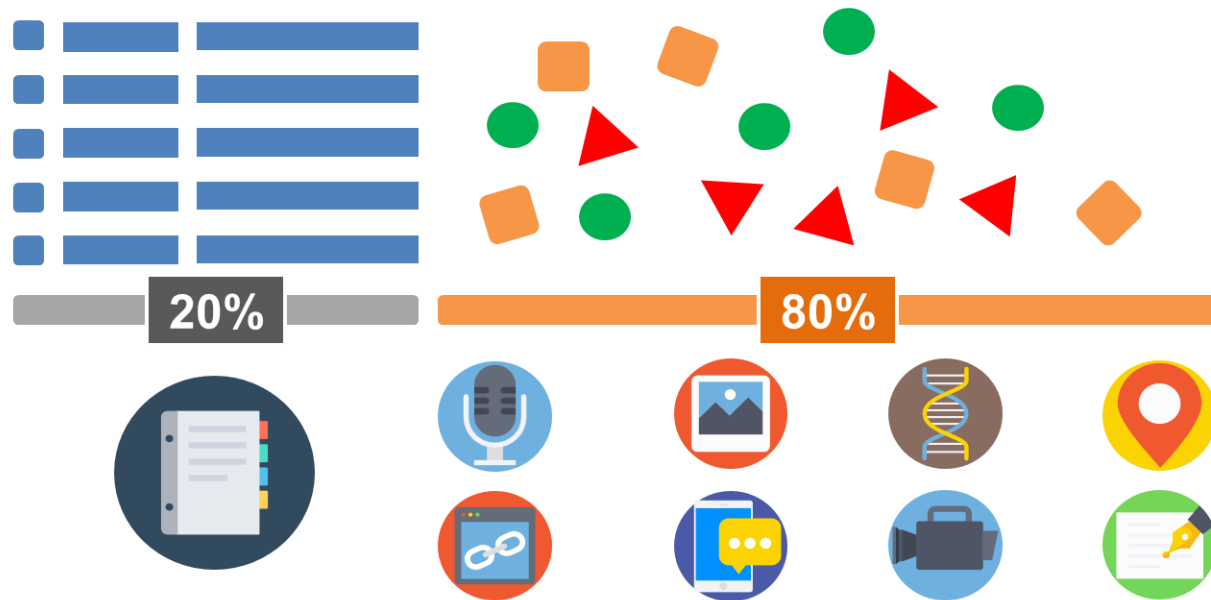
連續資料 (Continuous Data)

在任兩數值間可插入無限多數值之資料一般，
例如：身高、體重等



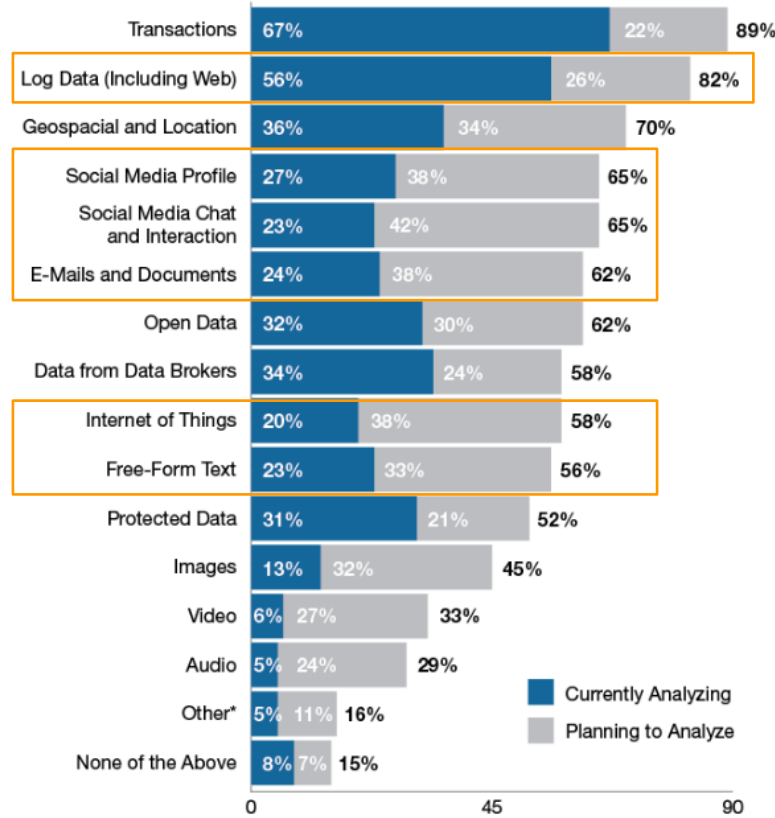
資料爆炸的時代中資料的儲存格式

- 20% 結構性資料，80% 半/非結構性資料



[href="https://www.flaticon.com/authors/prosymbols"](https://www.flaticon.com/authors/prosymbols)

Data Sources



n=195
 Source: Nick Haudecker, "Best Practices for Designing Your Data Lake," Gartner, 19 October 2016.
 *Other include genomic, meteorological, traffic, security, and customer data.
 © 2018 Gartner, Inc. and/or its affiliates. All rights reserved.
 Gartner is a registered trademark of Gartner, Inc. or its affiliates.

然後他就結構了



批踢踢實業坊 > 看板 Gossiping

看板 精華區 最舊 < 上頁 下頁 > 最新

搜尋文章...

- 4 [問卦] 有沒有義大利政局不穩的八卦?
IslamicState 8/21 ...
- 2 [問卦] 遇到盜攝集團該如何自保?
cychine 8/21 ...
- 4 [問卦] 小鳥被吹傷該怎麼辦? 在線等
ilikebrandy 8/21 ...
- [問卦] 肥宅我會不會練的太狀啊
malaman 8/21 ...
- [問卦] 請問擦了好癢 喉嚨痛的唇膏那裏買
wiwi0526 8/21 ...

PTTCrawl.mlx | Final_table = 13x4 t... — □ ×

| | Reply | Date | Author | Title |
|---|-------|--------|----------------|--------------|
| 1 | " 4 " | "8/21" | "IslamicState" | "[問卦] 有... |
| 2 | " 2 " | "8/21" | "cychine" | "[問卦] 遇... |
| 3 | " 6 " | "8/21" | "ilikebrandy" | "[問卦] 小... |
| 4 | " " | "8/21" | "malaman" | "[問卦] 肥... |
| 5 | " " | "8/21" | "wiwi0526" | "[問卦] 請... |
| 6 | " " | "8/21" | "go190214" | "[新聞] 公... |
| 7 | " " | "8/21" | "roanmaus" | "[問卦] 鮑... |
| 8 | " " | "8/21" | "Scion" | "Re: [新聞]... |
| 9 | " " | "8/21" | "bengowa" | "[問卦] 違... |

[半/非結構性資料]

非結構的純文字或是半結構的網頁原始碼

[結構性資料]

整理過後，欄位定義明確的表格資料

Data preprocessing with Live Editor tasks

The screenshot displays the MATLAB Live Editor interface for a task titled "Data Analytics - Load Forecasting Case Study". The task is "Load messy data", which involves loading data from "LETdata.mat" and inspecting it with "head(nyiso)".

The "Clean Missing Data" task is currently active. It shows the following configuration:

- Input data: nyiso
- Specify method: Moving median
- Cleaning method: Fill missing
- Method to fill missing entries: Linear interpolation

The results pane displays the output of the task, which is an 8x11 timetable:

| | Date | CAPITL | CENT |
|---|----------------|----------|-------|
| 1 | 05/01/2007 ... | 981.9000 | 1.571 |
| 2 | 05/01/2007 ... | 991.8000 | 1.568 |
| 3 | 05/01/2007 ... | 950.1000 | 1.560 |
| 4 | 05/01/2007 ... | 968.9000 | 1.560 |
| 5 | 05/01/2007 ... | 968.5000 | 1.555 |
| 6 | 05/01/2007 ... | 949.2000 | 1.564 |
| 7 | 05/01/2007 ... | 941.6000 | 1.538 |
| 8 | 05/01/2007 ... | 939.4000 | 1.557 |

Below the table is a line graph titled "Number of filled missing entries: 4". The graph shows the cleaned data (blue line) and the filled missing entries (red dots) over time from July 2007 to April 2008. The y-axis ranges from 200 to 1600.

Data preprocessing in Text

"Items are occasionally getting stuck in the scanner spools."

tokenizedDocument

10 tokens: Items are occasionally getting stuck in the scanner spools

removeStopWords

7 tokens: Items occasionally getting stuck scanner spools .

normalizeWords

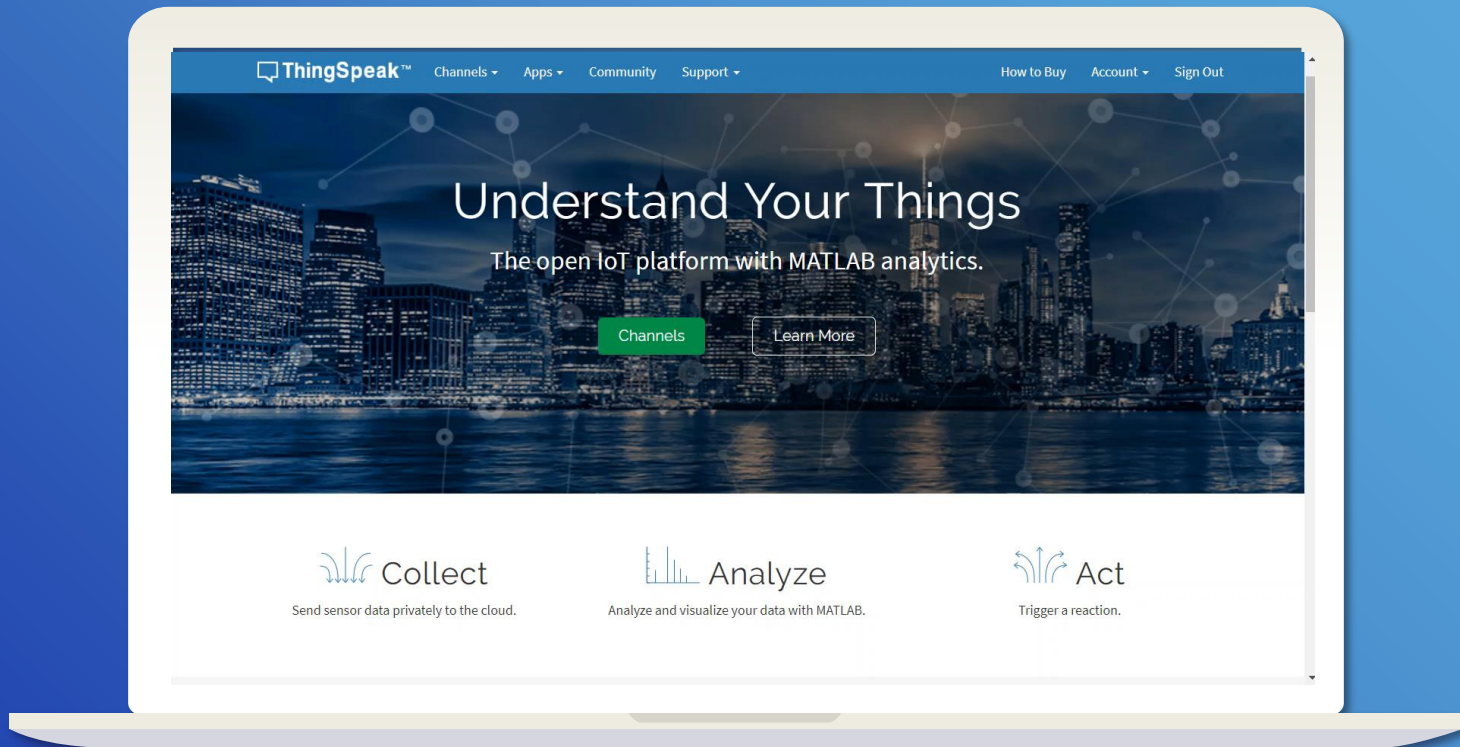
7 tokens: items occasionally get stuck scanner spool .

erasePunctuation

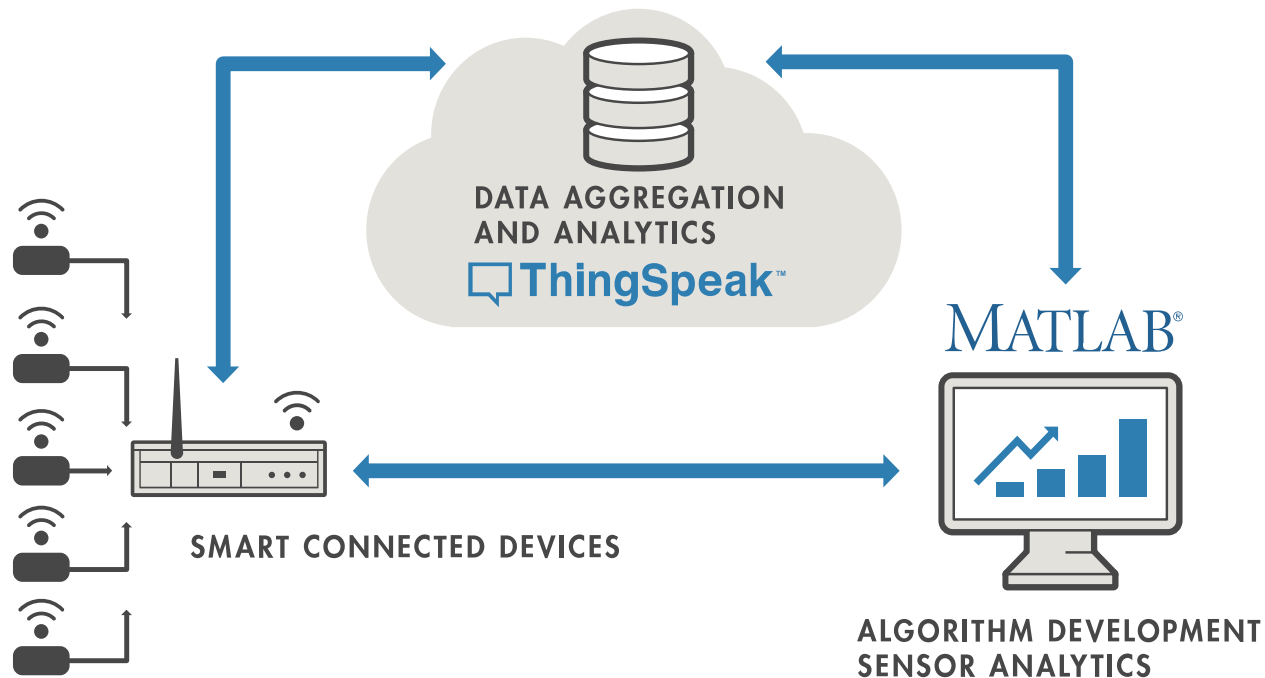
6 tokens: items occasionally get stuck scanner spool

雲端儲存平台-ThingSpeak

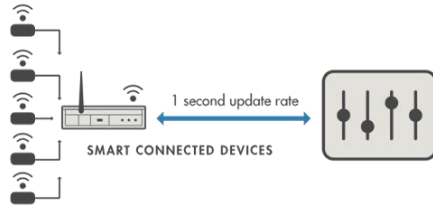
ThingSpeak



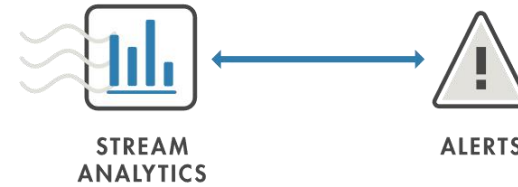
AIoT Environment in MATLAB



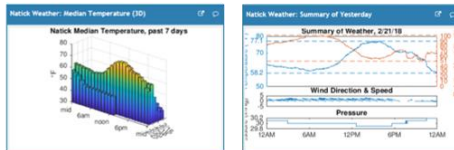
ThingSpeak Capability



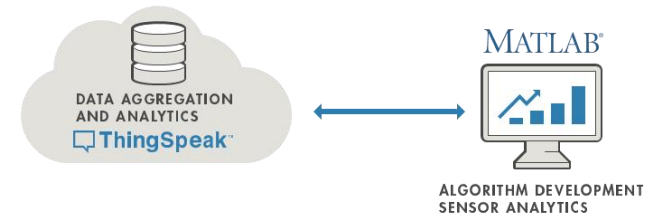
- 最高可更新速度 1筆/秒
- 可同時更新多個 Channel



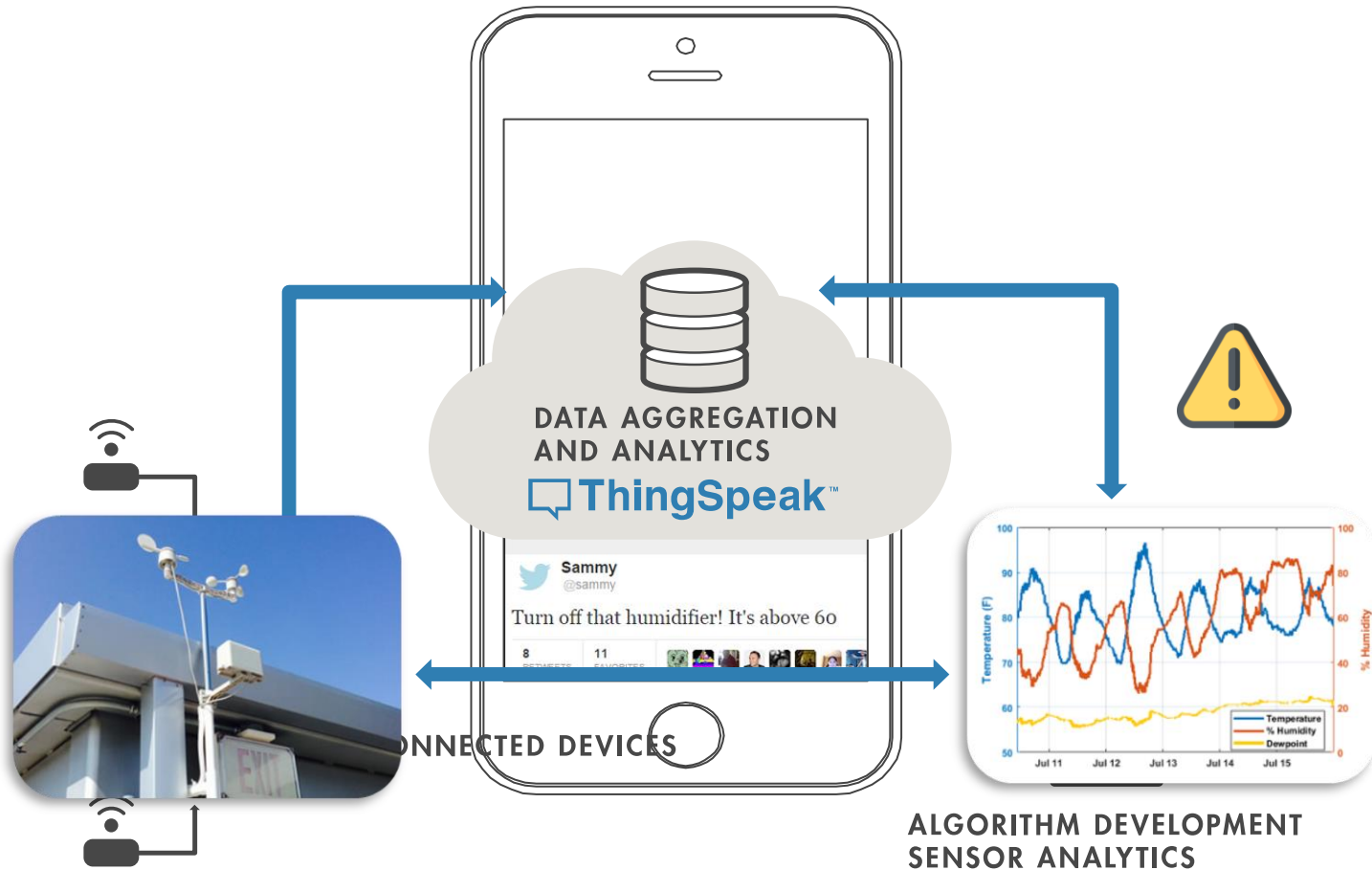
- 結合第三方應用程式推播



- 以MATLAB語法撰寫視覺化呈現

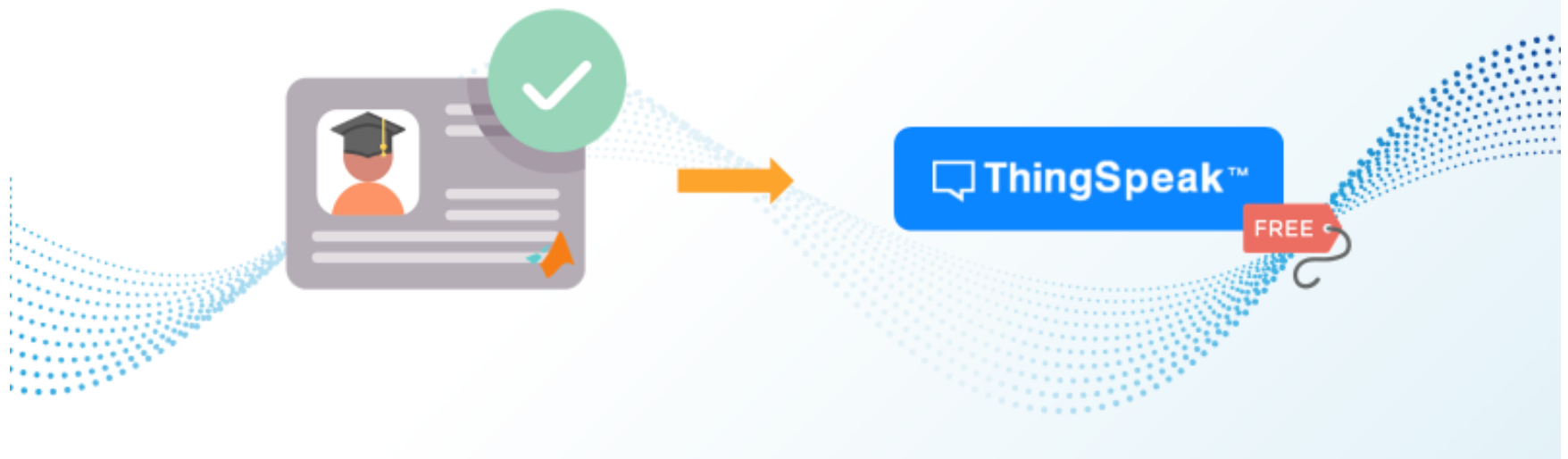


- 上傳或下載成 CSV · JSON 等檔案格式
- 在平台上以 MATLAB 撰寫資料分析演算法

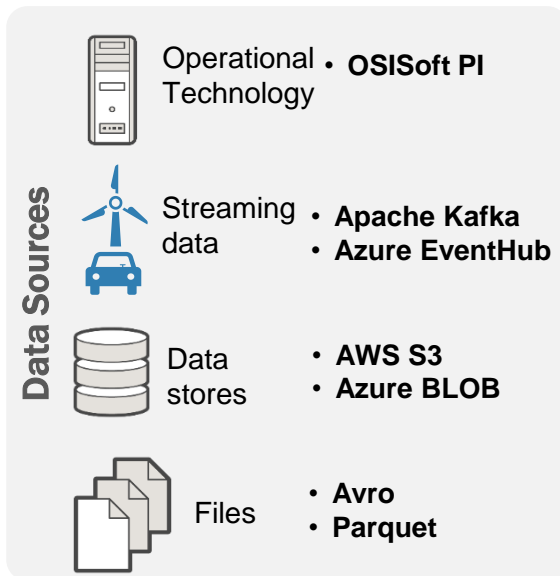


校園授權免費使用

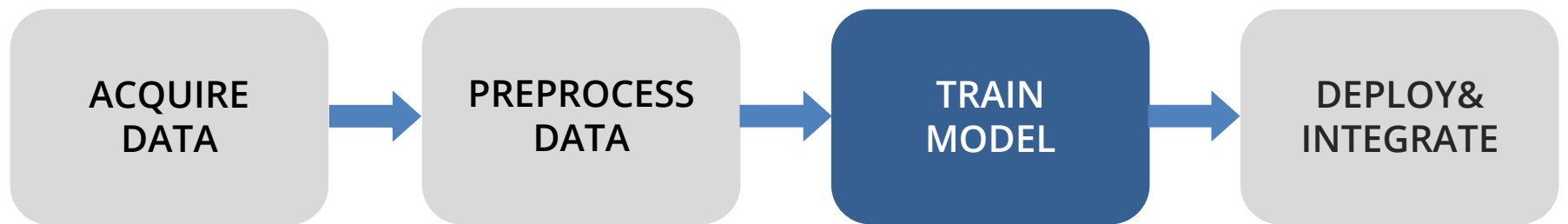
**Free Use with
Campus Wide License
(CWL)**



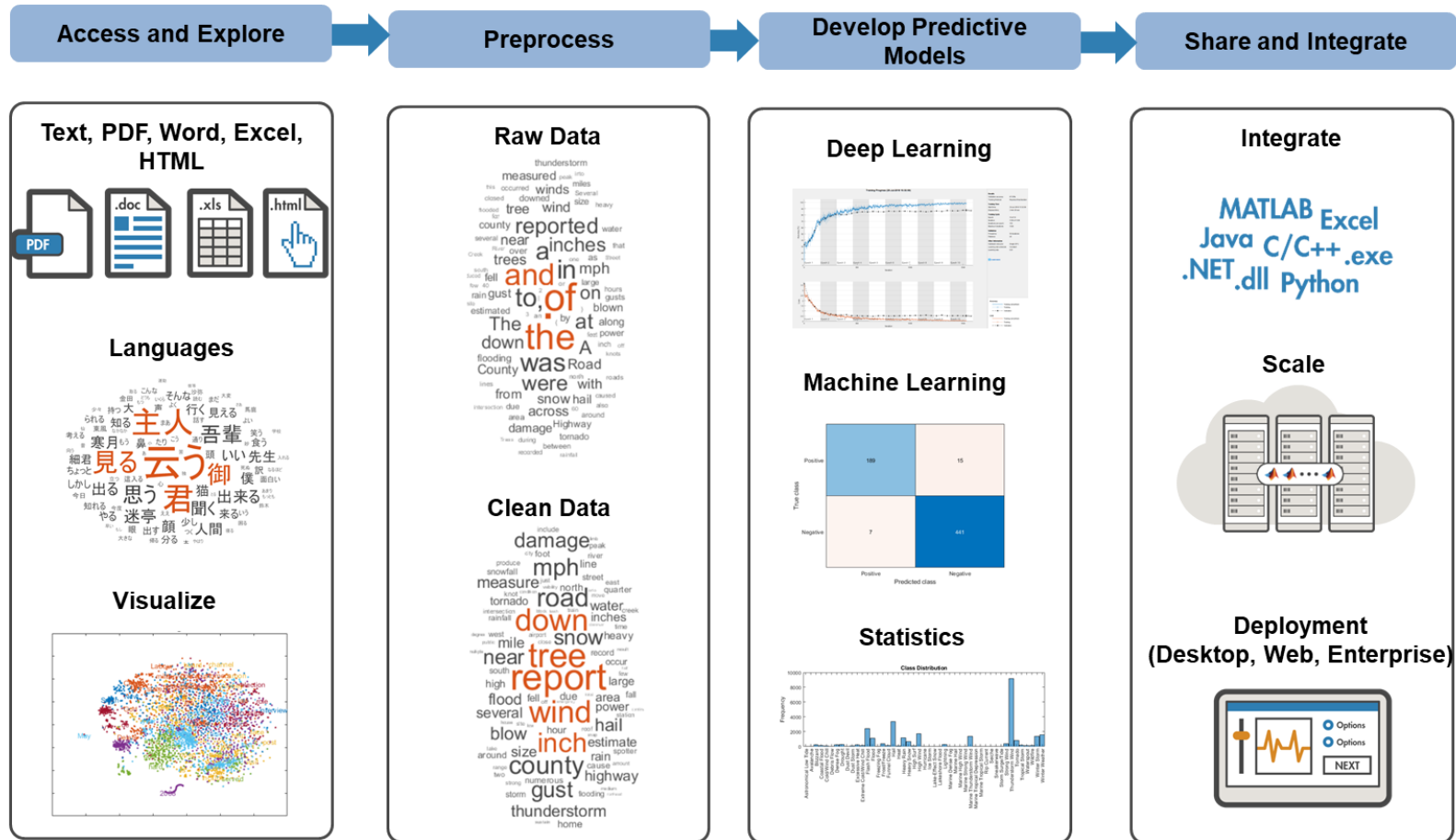
Public Cloud Storage



Scale to high-performance clusters using **MATLAB Parallel Server™**, and incorporate **MATLAB analytics** into applications with **MATLAB Production Server™**.

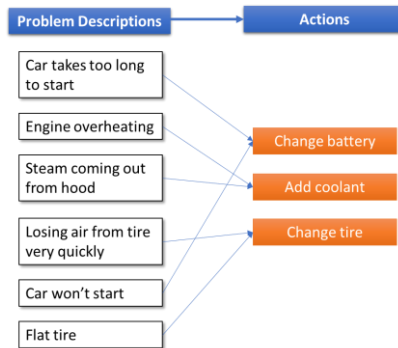


Text Analytics Workflow

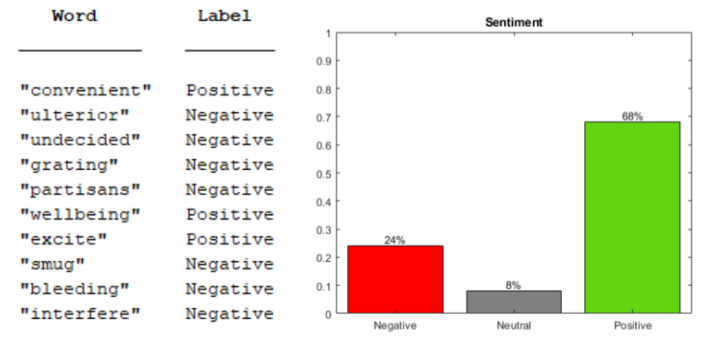


Text Analytics Applications

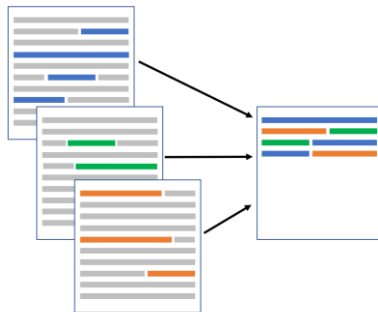
Classification



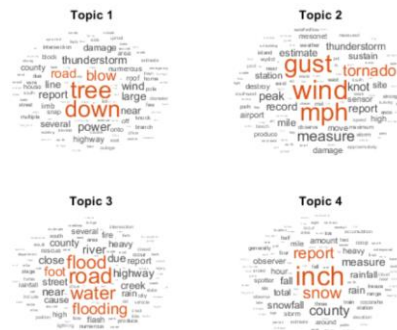
Sentiment Analysis

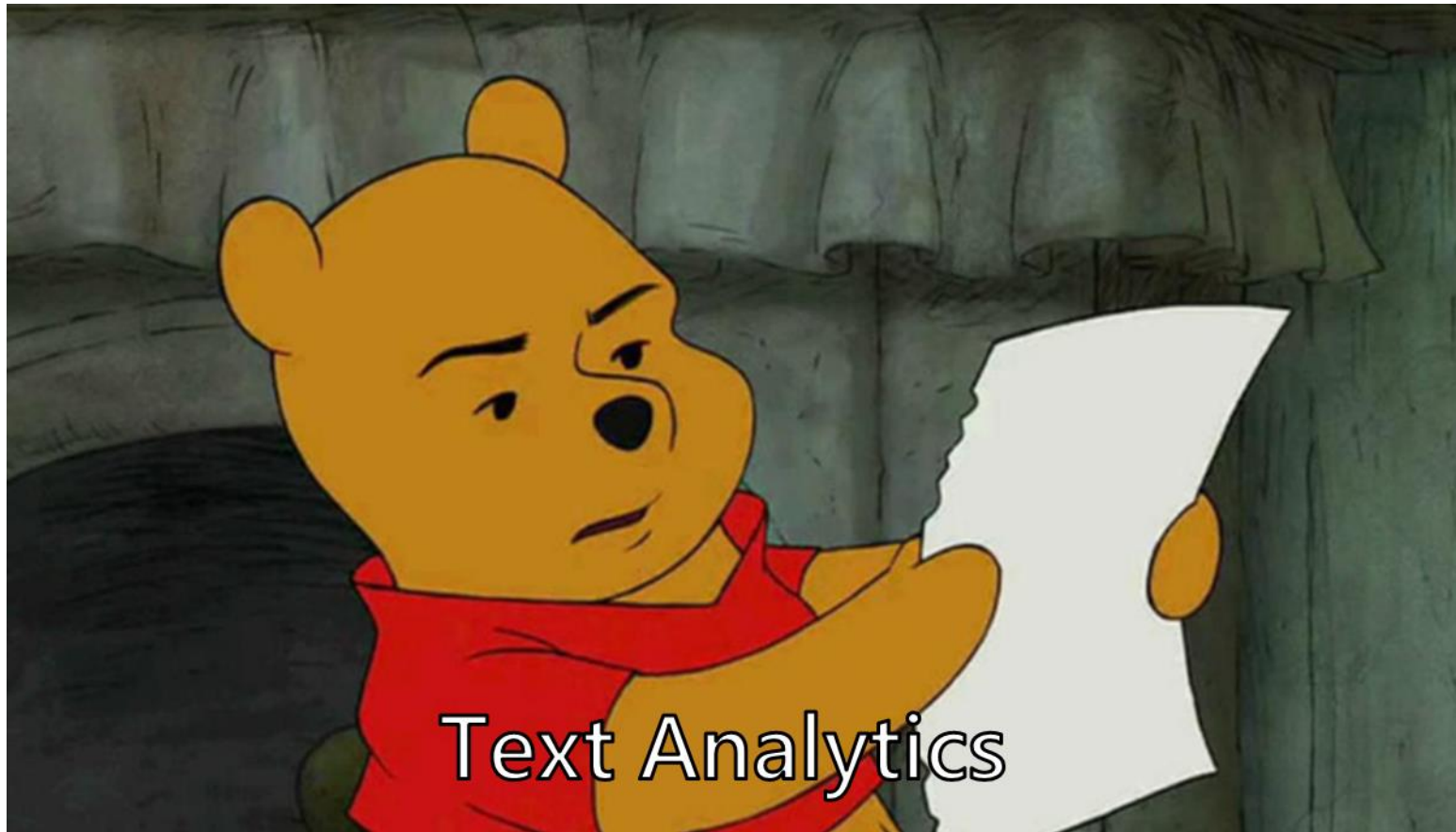


Summarization



Topic Modeling



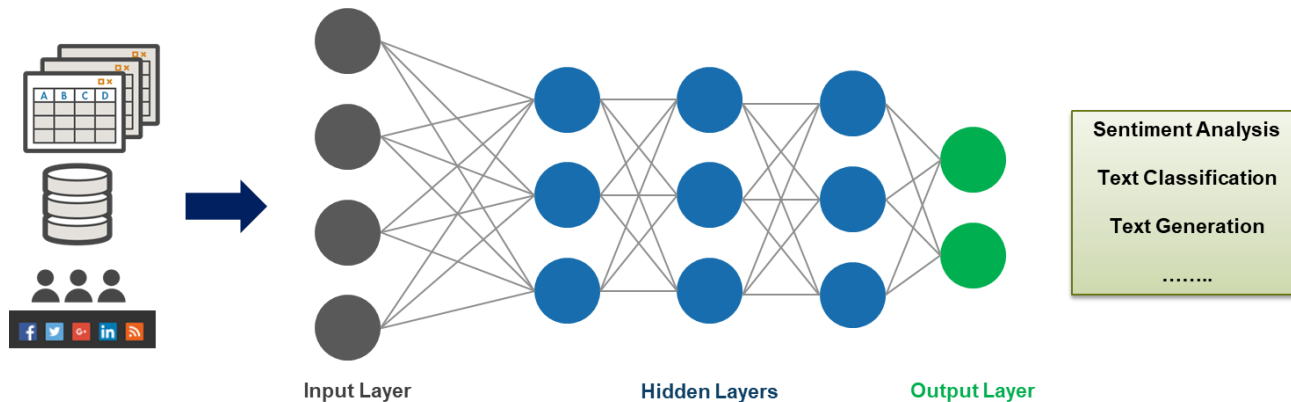


Text Analytics or Text Mining

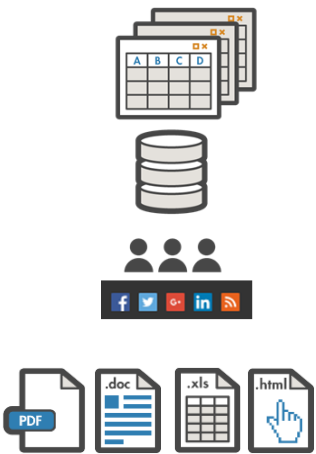


= Natural Language Processing + Machine Learning

Deep Learning and Text Analytics

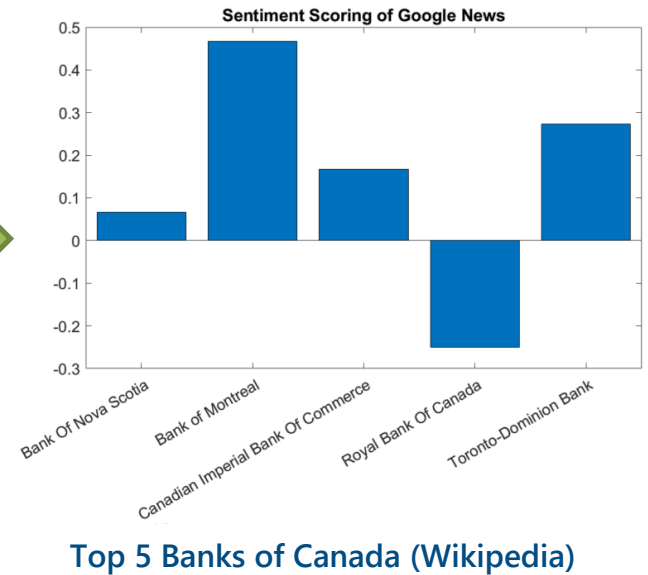


Train with APP



Deep Network Designer
app to build, visualize, and edit deep learning networks.

Classification Learner
App to train, visualize, and validation machine learning network.



Help | Modeling and Prediction — Examples

Documentation

Search Help

CONTENTS

- Documentation Home
- Examples
- Text Analytics Toolbox

Category

- Get Started with Text Analytics 6
- Toolbox
- Text Data Preparation 9
- Modeling and Prediction 23**
- Display and Presentation 2
- Language Support 2

All Examples Functions

Modeling and Prediction — Examples

Training Data

Classify Text Data Using Deep Learning

Classify text data using a deep learning long short-term memory (LSTM) network.

[Open Live Script](#)

Text Data: Preprocessed Bigrams

Analyze Text Data Using Multiword Phrases

Analyze text using n-gram frequency counts.

[Open Live Script](#)

Topic 1

Topic 2

Topic 3

Topic 4

Analyze Text Data Using Topic Models

Use the Latent Dirichlet Allocation (LDA) topic model to analyze text data.

[Open Live Script](#)

Raw Data

Cleaned Data

Create Simple Preprocessing Function

Create a function which cleans and preprocesses text data for analysis.

[Open Live Script](#)

Class Distribution

Create Simple Text Model for Classification

[Open Live Script](#)

Text Data: Preprocessed Bigrams

Analyze Text Data Using Multiword Phrases

[Open Live Script](#)

Topic 1

Topic 2

Topic 3

Topic 4

Analyze Text Data Using Topic Models

[Open Live Script](#)

Validation Accuracy

Choose Number of Topics for LDA Model

[Open Live Script](#)

The screenshot displays the GitHub interface for the repository `matlab-deep-learning/transformer-models`. At the top, there is a search bar and navigation links for Pull requests, Issues, Marketplace, and Explore. Below the repository name, there are tabs for Code, Issues, Pull requests (with a count of 1), Actions, Projects, Wiki, Security, and Insights. A dropdown menu shows the current branch as `master` with 2 other branches and 3 tags. A pull request by `bwdGitHub` is highlighted, showing a merge of pull request #10 from `matlab-deep-learning/fix_typo`. A list of files and folders is shown, including `+bert`, `+finbert`, `+gpt2`, `+sampling`, `+transformer/+layer`, `.circleci`, `test`, `.gitignore`, `ClassifyTextDataUsingBERT.m`, `FineTuneBERT.m`, `PredictMaskedTokensUsingBERT.m`, and `PredictMaskedTokensUsingFinBERT.m`. A repository overview card is overlaid on the right, showing the repository name, description, MATLAB logo, and statistics: 2 Contributors, 0 Issues, 42 Stars, and 13 Forks.

<https://github.com/matlab-deep-learning/transformer-models>

Demo case: Summarize from BBC news

C:\Users\tim\Desktop\ArticleSummary.mlx

LIVE EDITOR INSERT VIEW

FILE NAVIGATE TEXT CODE SECTION RUN

Summarize BBC News Using GPT-2 and Extract Keywords Using RAKE Model


NEWS

Home | Coronavirus | Video | World | Asia | UK | Business | Tech | Science | Stories | Entertainment & Arts

Business | Market Data | New Economy | New Tech Economy | Companies | Entrepreneurship | Technology | Economy | Global Car Industry

Cryptocurrency: Musk's SpaceX to launch dogecoin moon mission

3 days ago | Comments



從網頁讀取HTML純文字檔

```
url = 'https://www.bbc.com/news/business-57052811':
```

```
source =
16x1 htmlTree:

<P><B class="ssrcss-hmf8q1-BoldText e5tfeyi3">Elon Musk has said his rocket company SpaceX will now accep
<P>In a tweet he said the commercial space exploration firm will launch the "DOGE-1 Mission to the Moon"
<P>It came as dogecoin continued to fall after the technology billionaire said on television that it was
<P>Last month, he tweeted that SpaceX was going to put a "literal Dogecoin on the literal moon".</P>
<P>Canada's Geometric Energy Corporation announced the dogecoin-funded mission earlier on Sunday, although
<P>"This mission will demonstrate the application of cryptocurrency beyond Earth orbit and set the founda
<P lang="en" dir="ltr">SpaceX launching satellite Doge-1 to the moon next year<BR/><BR/>♦♦ Mission paid
<P>Mr Musk's tweets in recent months have helped to turn the once-obscure digital currency, which was sta
<P>It has soared by more than 800% over the last month, with its total market value rising above $70bn, a
<P>However, dogecoin lost more than a third of its price on Sunday, after Mr Musk called it a "hustle" wh
<P>In February, electric carmaker Tesla, where Mr Musk is chief executive, said it bought $1.5bn worth of
<P>The announcement was seen as another major step towards the mainstream acceptance of bitcoin and helpe
<P>Last month, Tesla announced that its profits for the first three months of the year were $438m, up fro
<P>Mr Musk also revealed on Saturday Night Live that <A href="https://www.bbc.co.uk/news/world-us-canada-
<P>People with Asperger's interpret the environment around them differently to other people.</P>
<P>It is thought to be the first time Mr Musk has spoken about his condition.</P>

str = 16x1 string
Rows 3:12
"It came as dogecoin continu...
>Last month, he tweeted that...
>Canada's Geometric Energy C...
>"This mission will demonstr...
>SpaceX launching satellite ...
>Mr Musk's tweets in recent ...
>It has soared by more than ...
>However, dogecoin lost more...
>In February, electric carma...
>The announcement was seen a...

inputText =
'Elon Musk has said his rocket company SpaceX will now accept the meme-inspired cryptocurrency dogecoin a

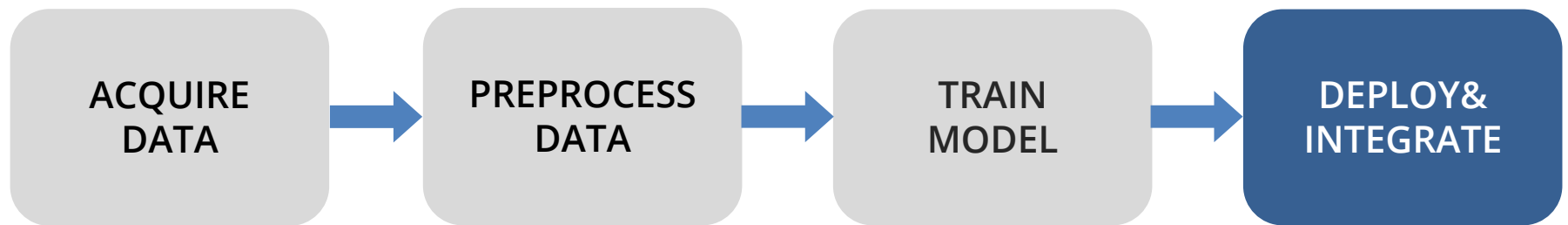
♦♦ Mission paid for in Doge
♦♦ 1st crypto in space
♦♦ 1st meme in space

To the moooooon!!! Mr Musk's tweets in recent months have helped to turn the once-obscure digital curren

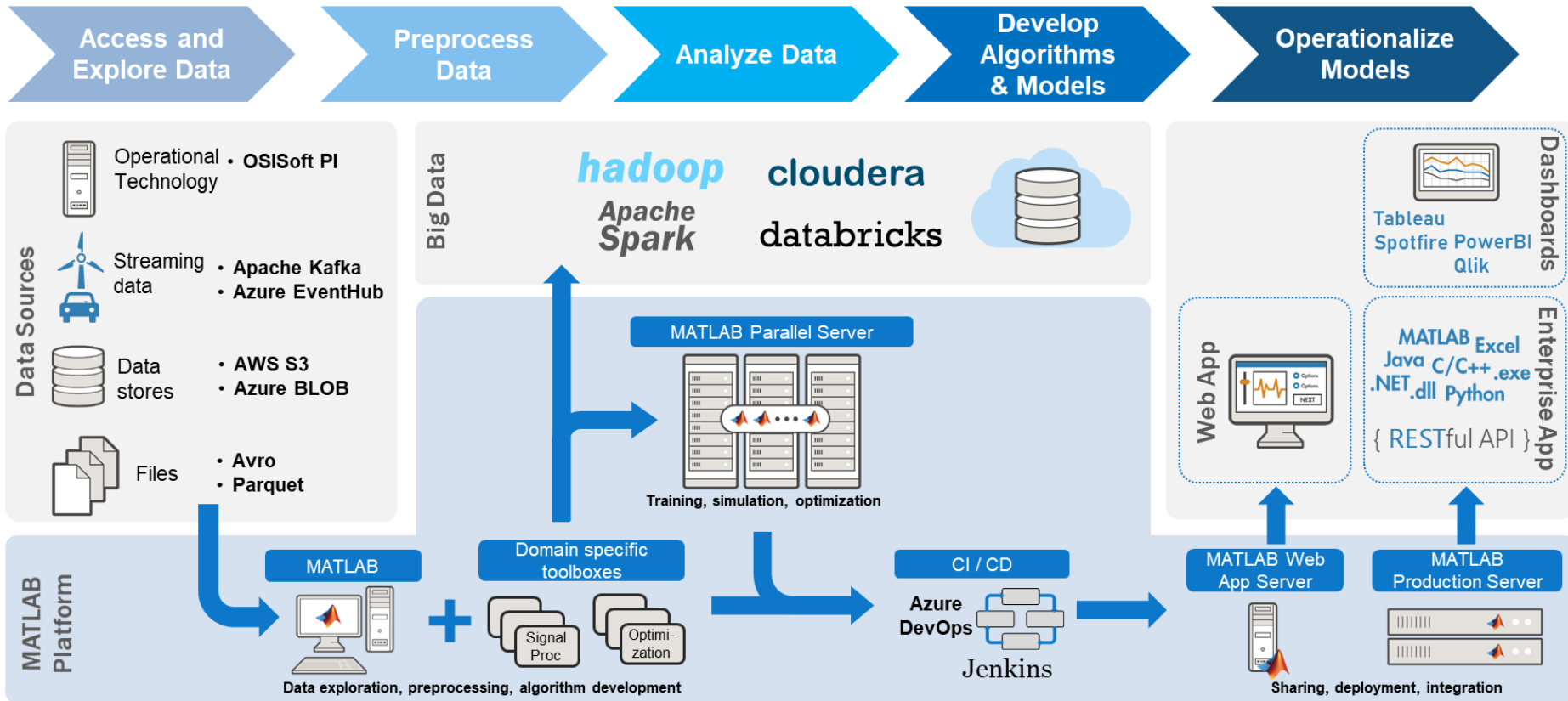
summary = ' Dogecoin, the cryptocurrency that has become the darling of the internet, is now accepted by the
```

UTF-8 LF script

Process workflow



MathWorks end-to-end solution



想提出問題?

您可以在演講進行過程中，利用問與答隨時向演講者問問題，您可透過工具列叫出問與答視窗。**請注意:**選擇“所有答疑者”詢問，答疑者了解問題並回答後便會顯示。

📌會議投影片: **06 /01 寄送至註冊信箱**

如何從工具列叫出問與答對話框提問?

1. 點選工具列中[...] 三個點選項，叫出更多功能



2. 從選單中點選問與答



3. 選擇“**所有答疑者**” 提問





TeraSoft 鈦思科技
www.terasoft.com.tw



MathWorks[®]