中華電信 Chunghwa Telecom

以 AI 之盾，防 AI 之矛

題目說明

近年來，一般使用者使用公開的服務，已有能力以生成式 AI 產生難以辨別真偽的語音與影像。本次將藉由各位同學的創意，應用生成式 AI 或 AI 偵偽/檢測技術，針對下列任一或多項議題，設計創新應用，以避免生成式 AI 合成語音及 AI 影像/影片生成對一般大眾的負面威脅。

本次挑戰議題：

1. 偽造合成語音偵測之技術及創新應用
2. 偽造影像生成偵測之技術及創新應用
3. Deepfake 假影音偵測之技術及創新應用
4. 例如應用於電信防詐、社群媒體防詐、假新聞防治與事實查核、社群媒體誤導/假資訊查核等（不限目前所列）

在定義問題並提出解決方案後，可利用公開之預訓練模型或線上 API（如 OpenAI/AWS/Google Cloud/Azure 提供之雲端人工智慧 API），搭配各類“NoCode”工具，快速搭建 MVP(Minimum Viable Product)方案。

評分標準

請參賽者現場 demo，並須附上**成果簡報**以利評分：

1. 成果簡報包含：

   目標對象、解決的問題、技術架構及服務期望值。

2. 評分面向：

   | 作品創意性：40%

   | 作品技術性：30%

   | 作品完成度：30%

中華電信 Chunghwa Telecom

相關資源

(一) 語音真偽：

語音偵偽 Mocking API (以 docker image 形式提供參賽者部署於自己的終端或雲端設備)

1. 公開模型與程式碼：https://paperswithcode.com/dataset/asvspoof-2019
2. ASVSpoof 競賽：https://www.asvspoof.org/
3. No Code 工具：https://marvelapp.com/

(二) AI 生成影像檢測：

AI 生成影像檢測 API (以 docker image 形式提供參賽者部署於自己的終端或雲端設備)

1. 公開模型與程式碼：

   | https://github.com/PeterWang512/CNNDetection

   | https://github.com/WisconsinAIVision/UniversalFakeDetect

2. 公開資料集：

   | https://github.com/PeterWang512/CNNDetection

   | https://www.kaggle.com/datasets/awsaf49/artifact-dataset