FORTINET | 2025 Fortinet 資安嘉年華
Resilience Powered by AI and Cybersecurity

最後的防線：
用 AI 對抗 AI，守護你的新世代工作空間

Josh Lin 林裕祥
Fortinet 首席工程師

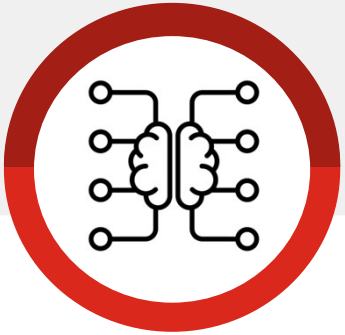
AI 發展的 4 個階段：從助攻生產力到進化攻擊力

Shadow AI
員工私自使用 AI

- 員工自行使用 ChatGPT、Gemini、Grok 等外部 AI
- 用戶端/瀏覽器層級的使用，企業沒有可視性

SaaS AI
SaaS 內嵌 AI

- 使用的 SaaS 工具開始加入內建 AI 功能
- Microsoft 365 Copilot、Google Workspace Gemini，自動讀取更多文件/郵件



Private AI
企業自建 AI

- 企業開始把內部資料導入 AI，建立私有模型，AI 回答問題時會先查資料，然後再生成答案
- 大部分企業使用 RAG，外掛公司資料庫

Offensive AI
AI 被駭客利用

- AI 變成攻擊自動化武器，讓初階駭客也能做高階攻擊，破解安全政策自動繞過防禦
- 攻擊從「人力密集」變成「AI 自動化」

FraudGPT × Deepfake × 仿聲詐騙，全面升級攻擊

AI 已變成駭客的外掛，生成式攻擊正在快速氾濫

FraudGPT ElevenLabs 、BlackmailerV3 等工具可自動化產生惡意程式、釣魚網站、深偽影片以及仿聲詐騙 → 讓攻擊更具規模、更可信、更有效

57%

社交工程攻擊事件中有 57% 與釣魚相關

$2.8B

2024 年 (僅美國) 因 BEC (商業郵件詐騙) 攻擊造成的損失

70%

過去 12 個月內，七成企業成為 BEC 攻擊的目標

44%

勒索軟體的入侵事件 (去年為 32%)

42%

暗網上出售的被竊憑證數量增長

Web 世界危機四伏！ 網頁早已成為駭客的「百寶袋」

Phishing / 網路釣魚
Malware and Ransomware / 惡意與勒索軟體
SQL Injection / SQL 注入攻擊
Browser Hijacking / 瀏覽器劫持
Click-jacking / 點擊劫持
Cross-site Scripting / 跨腳本攻擊
Drive-by Downloads / 偷渡式下載
Form-jacking / 表單劫持
Man-in-the-Middle Attacks / 中間人攻擊
Session Hijacking / 會話劫持
Spyware / 間諜軟體
DNS-based Threats / DNS 威脅
Malicious URLs / 惡意網址

攻擊面大爆發！無處不在的攻擊入口

駭客鎖定的痛點

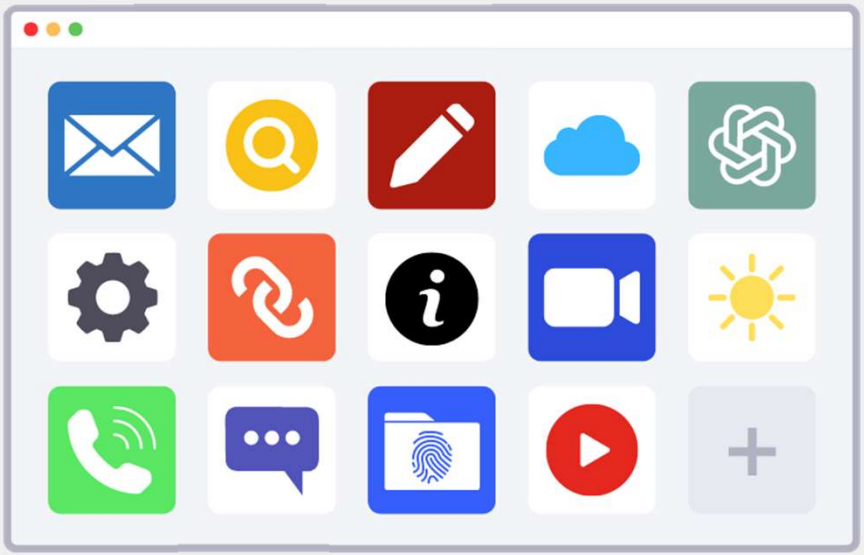
郵件、雲端、協作工具，全面淪陷

網路釣魚

攻擊者用假郵件或訊息引誘點擊惡意連結

內部惡意人員

公司內部人員蓄意洩漏或竊取資料



惡意軟體下載

使用者不小心從協作平台、郵件附件或網頁下載木馬病毒

高風險擴充功能

瀏覽器或應用程式外掛被駭客利用，變成後門

第三方風險

外部合作夥伴或供應鏈帳號被入侵，成為攻擊跳板

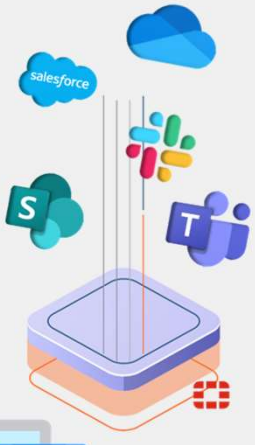
新世代工作空間框架

電子郵件防護

網頁瀏覽器

零信任存取控制

SaaS 雲端應用



協作平台

端點防護

協作平台 = 駭客的狩獵場!

協作平台已成為重點攻擊場域

網路釣魚

帳號接管

偽裝身分

惡意軟體

勒索軟體

內部人員風險

社交工程

我需要客戶銷售數據來做分析

@

我分享一篇文章給你，看一下

@

Carl 要我找你拿新產品的設計檔

可以把測試結果分享給我嗎

@

我們需要在今天結束前付款給供應商，才能拿到早鳥優惠

新世代工作已轉型，但你還停留在單點防禦？

頭痛醫頭，腳痛醫腳

當你還在局部止痛
駭客早已透視你的全身!

廣告信過濾


機敏內容過濾


寄件者位址比對

SPF 驗證

URL 類別過濾


黑白名單比對


靜態連結檢查

AI 驅動 × 全方位防護，瀏覽上網 、郵件與協作零死角

提供 AI 驅動的企業級防護，可依需求靈活客製化

- 防護進階釣魚、BEC (商業郵件詐騙) 、冒充、帳號接管與勒索攻擊
- 企業級功能，適用於各種規模組織與應用情境
- 提供多種部署模式，包括 : 安全郵件閘道器 (SEG) 與整合式雲端郵件安全 (ICES)
- 超越郵件防護範圍，涵蓋安全瀏覽與協作平台，結合專家 24/7 支援

協作平台安全 | 瀏覽器安全 | M365 整合 | Google 整合

全通路內容防護

高度擴展性與靈活性的解決方案，即便是最複雜的環境也能輕鬆應對

AI 驅動 × 內容智能分析，守護機敏資料零死角

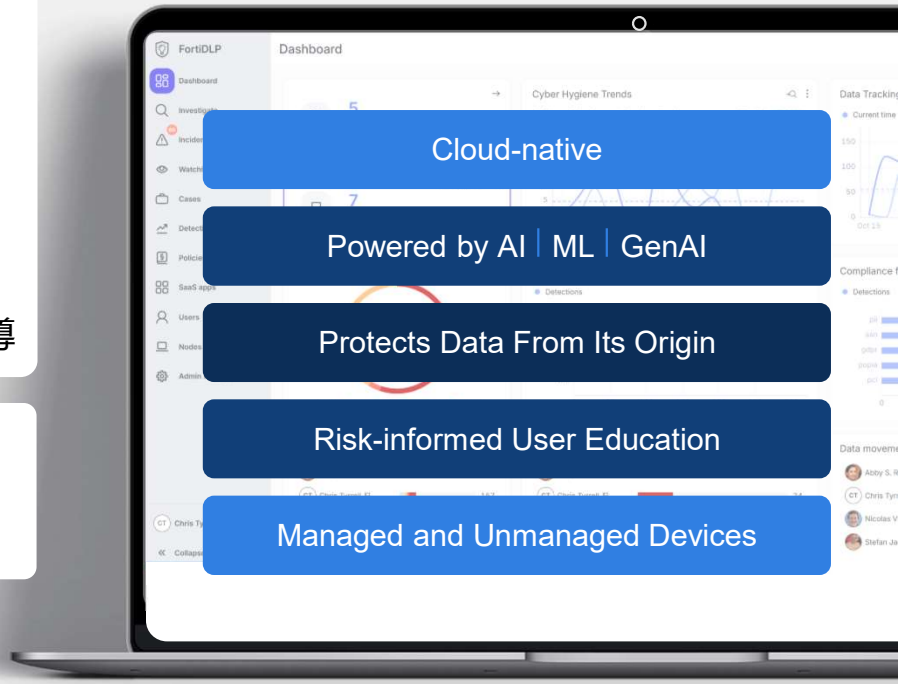
新世代DLP 和內部風險管理解決方案，可預測並防止數據竊取

- 對資料流和風險的全面可見性
- 即時上下文和內容等級檢查
- 跨 SaaS 應用程式的資料保護
- 專注於資料訪問時，根據風險對用戶進行宣導

資料外洩防護

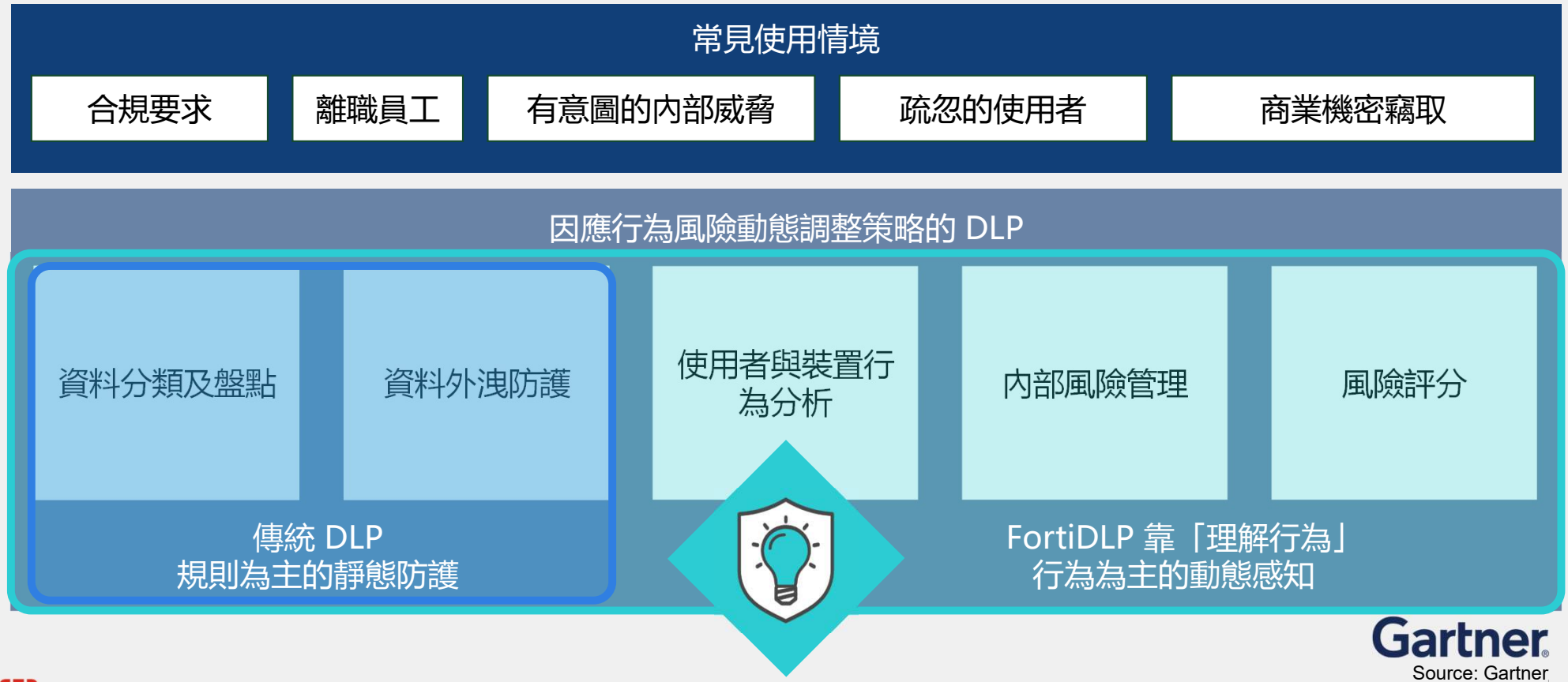
內部風險管理

SaaS
數據安全



自動將檢測映射到 MITRE ENGENUITY™ 內部威脅 TTP 知識庫

AI 驅動攻擊來了，還在用關鍵字 DLP 擋資料外洩嗎?



Gartner

Source: Gartner

情境分享：DNS，最安靜的外洩通道

情境說明

公司有 HTTP/HTTPS Proxy 與上傳管制， 員工下班後到星巴克改用 DNS 查詢通道 傳資料：

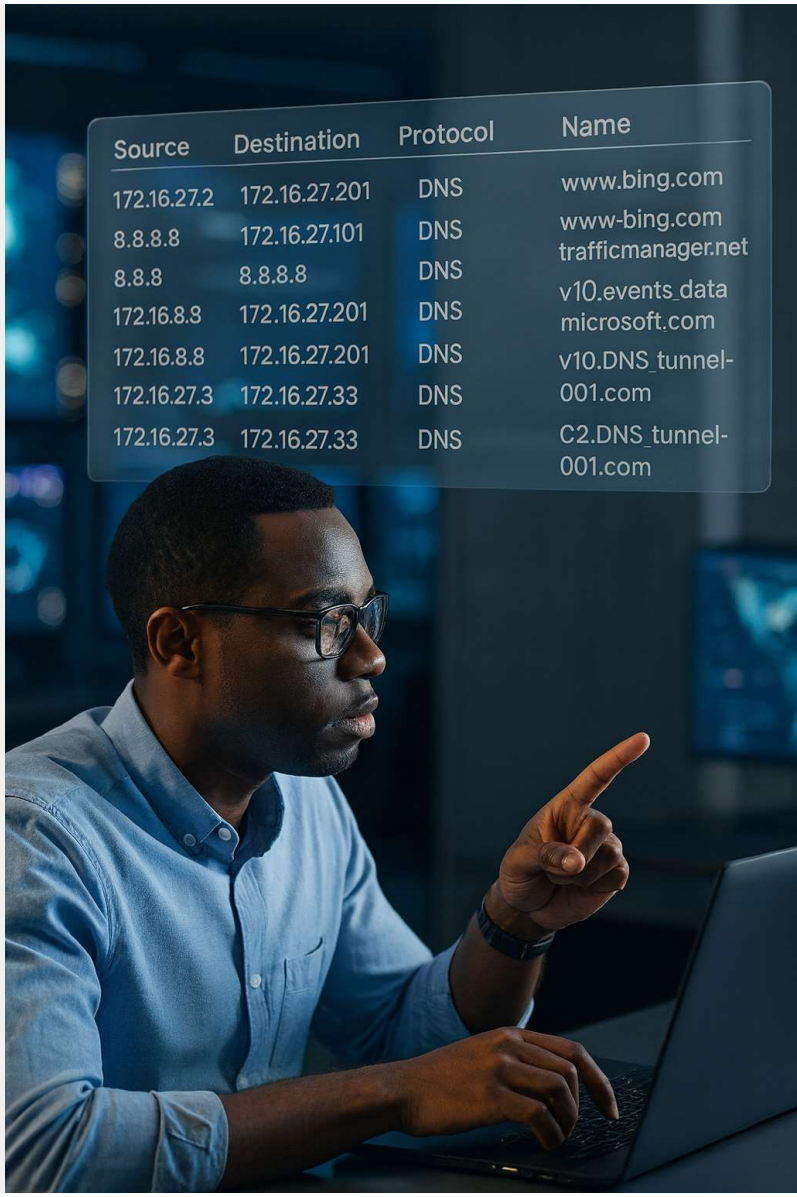
- 把資料切成一小段一小段，編碼後塞進像 **abc123.customer-db.backup.example-attacker.com** 這種奇怪的子網域
- 有心的使用者可以透過 DNS 工具與外部 DNS Server 建立連線，把資料切成多段查詢送出，並在外部伺服器將每一段資料重新拼回完整內容
- 對一般監控來說，看起來只是「很多失敗的 DNS 查詢」，但其實是機敏資料慢慢被搬出去

傳統 DLP 會怎麼做

多半只盯著「檔案上傳」這類通道：檢查寄出的郵件附件、HTTP/HTTPS 上傳或複製到 USB 的動作，卻不會去重組或分析 DNS 封包本身的內容在傳統監控畫面裡只像是一台電腦頻繁查詢怪異網域

AI DLP 會怎麼做

不只掃「檔案」，還會結合網路與使用者行為分析。針對 DNS 非典型通道，會偵測主機是否突然產生大量、異常的編碼查詢，並關聯這台主機上的機敏資料存取紀錄，判定為可疑外洩時直接告警或阻擋，不是只在有檔案上傳時才動作



情境分享：列印量暴衝， 內部作業還是威脅?

情境說明

某財務主管負責結算季報，手上有完整的營運數字與客戶應收款資料，平常只會列印幾十頁對帳單。但某一天在下班前，一台非公司管控、私自安裝的印表機突然收到幾百頁的大量列印工作，文件內容全是完整的客戶清單與財務報表

傳統 DLP 會怎麼做

只能做到「能不能列印」或「加不加浮水印」，它看的是「這份文件是否含機敏內容 → 允許列印 / 阻擋 / 加浮水印」，但缺乏對「誰在什麼時間點，突然大量列印異常頁數」這類行為的分析與關聯

AI DLP 會怎麼做

不只看「有沒有列印」， 還會把列印的使用情境一起納入判斷
監控所有列印動作（包含本機、USB、Wi-Fi 等私接印表機），對被標記為機敏的文件， 也會參考這個使用者平常是否會列印這類文件、這台裝置是否常連這台印表機、列印時間是否異常、列印頁數是否明顯偏高

情境分享：機密資料在 SaaS 之間偷偷搬家

情境說明

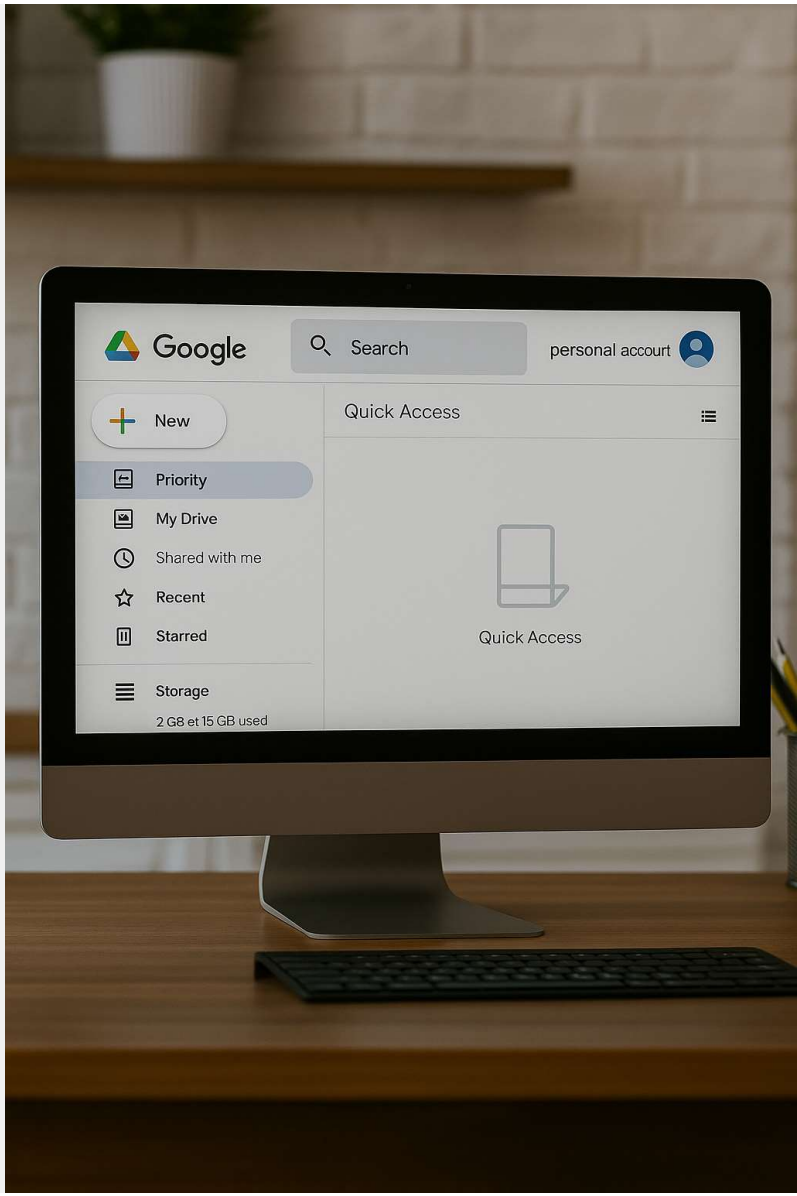
員工平常只是正常編輯文件，並且保存在公司 OneDrive。某天他將含有完整客戶清單的檔案先從公司雲端下載到自己電腦，再複製一份、改檔名，最後上傳到個人 Google Drive 或其他未核可雲端空間

傳統 DLP 會怎麼做

多半分別看「從 OneDrive 下載」與「上傳到 Google Drive」這兩個動作， 很難串起整條從企業雲端 → 本機 → 個人雲端的資料移動路徑， 不會知道這是公司授權的 OneDrive 租戶，還是員工個人的雲端空間，對於各種 SaaS 應用程式也缺乏細緻識別能力

AI DLP 會怎麼做

資料從下載到端點就會被標記是否為敏感資料，所有操作在端點都會被追蹤，一旦文件被發送出去，會追蹤數據來源。無論上傳的檔案是否可以讀取，它都會比較來源是否標記為敏感資料，如果來源敏感，立即阻止上傳操作, 並保留完整「從企業雲端 → 端點 → 私人雲端空間」的追蹤路徑



雲端郵件安全 × 釣魚防禦專家

防禦所有 AI 偵測的釣魚攻擊手法

兩步驟釣魚偵測

物件偵測模型會檢查網頁，辨識可點擊的元素以進行進一步掃描（常見的規避手法：使用者首先會看到一個乾淨的頁面，當點擊其中的元素後，才會被重新導向至隱藏的陷阱頁面，如釣魚頁面或惡意軟體下載）

相似網域偵測 (Domain Lookalike Detection)

S-GLocal 演算法結合啟發式生物演算法 (如 SLAGAN、區域比對與全域比對)，並針對電子郵件威脅中特別調整，用來精準識別相似網域的仿冒攻擊； 就像是 DNA 比對一樣，用演算法去比對網址的『相似度』

登入表單偵測 (Login Forms Detection)

影像辨識模型可偵測輸入框與登入表單，與 URL 進行交叉比對，找出異常以防止憑證竊取

品牌辨識

AI 會比對郵件、網址、截圖、圖片與小圖示 (favicon)，檢查是否假冒知名品牌。系統會同時參照官方的「正版圖像」 (如 Microsoft 商標) 以及已知的「惡意圖像」(例如仿冒樣本)，快速識別釣魚與仿冒攻擊

URL 語法分析 (URL Lexical Analysis)

AI 會分析 URL 的結構，找出與惡意 URL 的相似性，並預測該 URL 是否具有惡意。搭配 OCR (文字辨識)，從圖片 / PDF 中抽取並分析網址，以偵測潛在威脅

異常偵測 (Anomaly Detection)

偵測寄件者語氣與情緒的變化，分析郵件內容並與中繼資料進行比對，以發現異常偏差；同時利用主題建模 (topic modelling) 分析郵件主旨與內文，找出更多可疑跡象

雲端郵件安全 × BEC 攻擊剋星

AI 抵禦所有 BEC 威脅手法

GenAI Decoder

LLM 模型能辨識 AI 生成文字的特徵，揪出詐騙郵件與社交工程攻擊

郵件討論串劫持防護 (Thread Hijack Protection)

透過網域仿冒關聯演算法，防止郵件對話、回覆被劫持，以及供應商冒充攻擊

內容分析 (Content Analysis)

多種 NLP 模型能抽取敏感內容 (如個人可識別資訊 PII)、識別實體名稱 (NER)、分析文字特徵的附加資訊，像主旨、字數長度、用詞風格，AI 會用這些線索來判斷這封信正不正常

供應鏈識別 (Supply-Chain Recognition)

分析業務溝通，自動識別合作夥伴、供應商等相關網域，避免駭客假冒

異常偵測 (Anomaly Detection)

偵測寄件者語氣與情緒的變化，分析郵件內容並與中繼資料進行比對，以發現異常偏差；同時利用主題建模 (topic modelling) 分析郵件主旨與內文，找出更多可疑跡象

文字混淆偵測 (Textual-Obfuscation Detection)

偵測駭客透過字元替換或隱藏來掩蓋惡意文字的規避手法，例如用外觀相似的字元替換 (O → 0)，或插入不可見字元等

AI × 圖像雙重防禦，釣魚詐騙無所遁形

GenAI Decoder

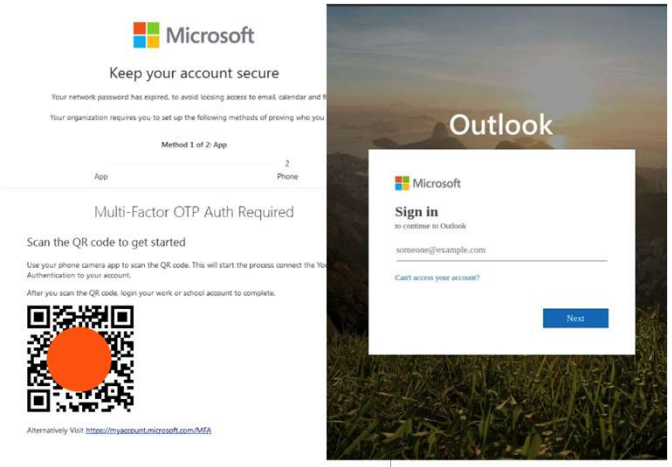
偵測魚叉式網路釣魚 (Spear Phishing)、BEC / VEC（商業 / 供應鏈郵件詐騙）
使用 LLM（大型語言模型），專門辨識 AI 生成文字, 判斷一段文字是不是由 AI 生成的

Confidential surprise	I must reward the commitment to work and the dedication of the staff during this period. Therefore, I am **Surprising some executive staff** today, and I would appreciate your confidentiality so as **not to spoil the impact of the surprise.**
Call to action	I will need you to make an online purchase on my behalf or at any stores near you for gift cards to say thank you, and well done. Any such cards like **Amex /Visa or Amazon gift cards** for variety depending on their tastes. **Your input on this idea** would be appreciated before purchasing, so I know what you as a staff feel about this;
Urgency	My schedule is quite tight during this period, but please reply to the email as soon as you get this so we can find the best possible way to do it.

Image Recognition

偵測 Quishing (QR Code 釣魚)、一般釣魚網站

透過圖像辨識，解析圖片內容，並比對檔案截圖、URL 圖示 (favicon) 是否冒充知名品牌



新世代瀏覽器安全，讓你的瀏覽器變身安全堡壘

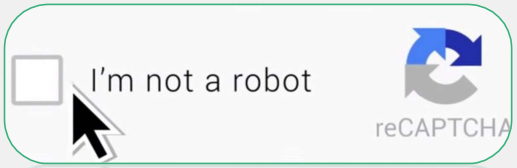
在標準瀏覽器中加入 AI 偵測、資料外洩防護 (DLP)、瀏覽器級別治理， 在瀏覽器裡加上安全規則和控管，讓每一個上網動作都受到公司安全政策的保護

提供動態防護，對使用者體驗與瀏覽品質零影響

全天候 24/7 託管事件回應，透過跨通訊與應用的情境關聯分析，快速拼湊出威脅全貌，偵測一般防護難以發現的隱藏攻擊

透過輕量級瀏覽器擴充套件，即時集中部署，支援 Chrome / Edge / Safari / Firefox 等

瀏覽器防護新境界，守護你的辦公日常



新世代網址反規避技術 (Next-Gen URL Anti-Evasion)

可偵測釣魚攻擊，即使駭客使用 CAPTCHA 、人為地理位置限制 (Geofencing)、或時間型規避手法也能識別

瀏覽器內建資料防護 (Browser-Based DLP)

可偵測使用者在任何網頁應用 (如 GenAI 、電子郵件等) 中上傳或下載敏感內容

使用者安全意識 (User Security Awareness)

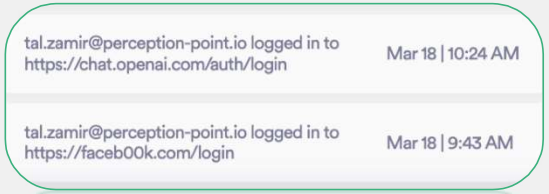
警示使用者避免不良習慣，例如密碼重複使用，或在第三方網站可能導致資料外洩的風險

關聯瀏覽器與電子郵件事件 (Correlating Browser & Email Events)

在瀏覽器中開啟的URL，會自動關聯回來源郵件，方便調查追蹤

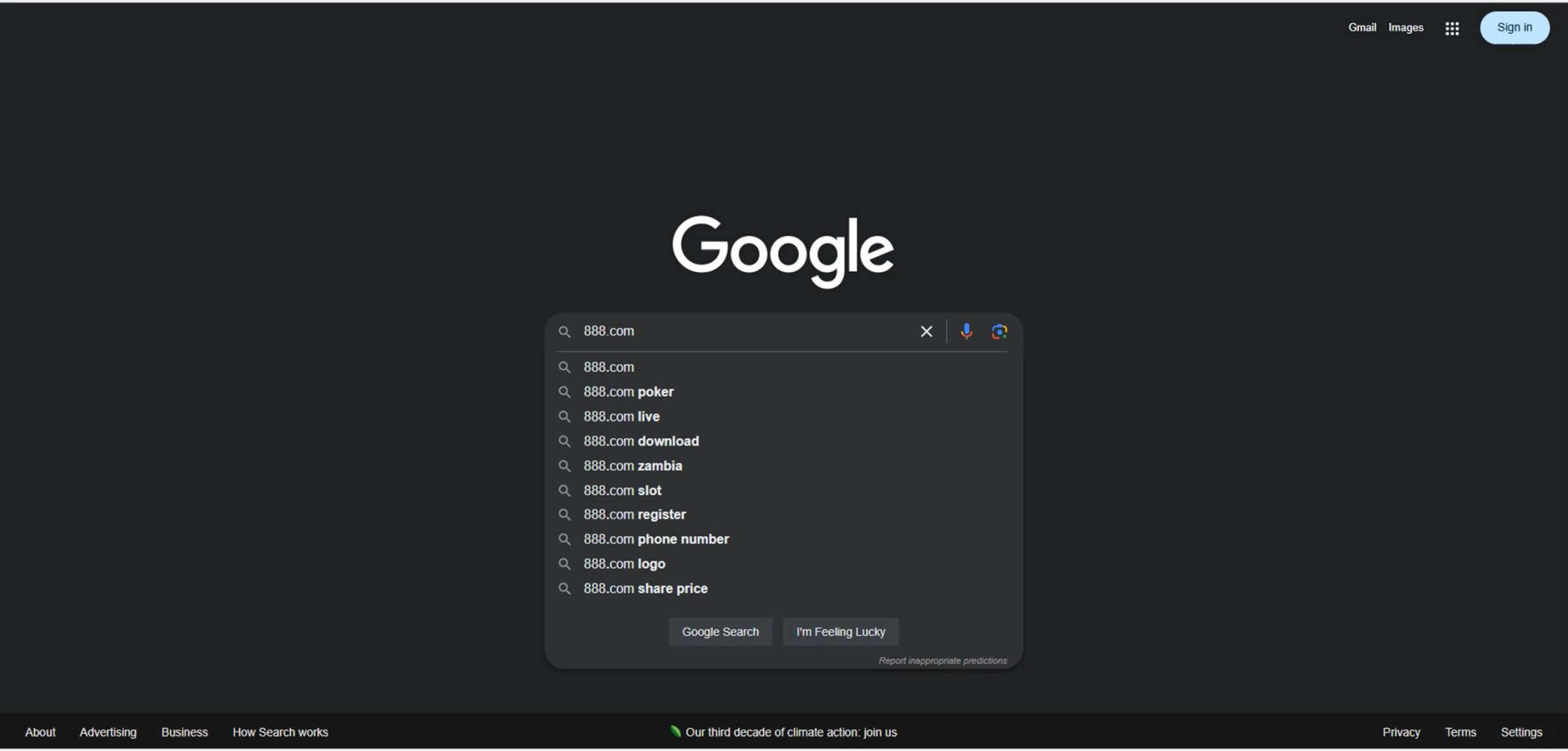
跨應用自動隔離 (Cross-App Remediation)

可同時隔離受威脅的郵件或檔案，例如瀏覽器偵測到惡意網址時，會自動隔離所有相關郵件



入侵後影響分析 (Post-Breach Impact Analysis)

可掌握使用者登入事件，識別遭受釣魚的帳號，並執行補救行動


Gmail
Images
Sign in
Google
888.com
888.com
888.com poker
888.com live
888.com download
888.com zambia
888.com slot
888.com register
888.com phone number
888.com logo
888.com share price
Google Search
I'm Feeling Lucky
Report inappropriate predictions
About
Advertising
Business
How Search works
Our third decade of climate action: join us
Privacy
Terms
Settings

FORTINET

2025 Fortinet 資安嘉年華

Resilience Powered by AI and Cybersecurity