「護數守設 安全之道」

守護 AI 時代下的香港 K12 教育生態

網絡安全、數據私隱
合乎道德的AI應用

Ken Chung
K-Solve Global

K12 教育中的 AI 網絡安全與數據安全

教育領域的AI應用：機遇與風險

香港K12機構面臨的主要威脅

保障AI系統安全的最佳實踐

案例研究與經驗總結

CYBER EASY PLATFORM

K12教育中的 AI 網絡安全與數據安全簡介

K12教育中的 AI 網絡安全與數據安全簡介

隨著人工智慧（AI）技術在教育領域的快速普及，
K12教育機構正經歷數位化轉型的浪潮。
從智慧教室到個性化學習平臺，AI應用顯著提升了教學效率與學生體驗。
技術的深度整合也伴隨著新型網絡安全威脅與數據隱私風險。
尤其是K12教育機構，因其服務對象為未成年人，
且涉及大量敏感個人數據（如學生身份、學習記錄、家庭資訊），
使其成為網絡攻擊的高風險目標。如何平衡AI技術的創新應用與安全防護，
已成為全球教育機構的關鍵挑戰。
探討AI在教育中的雙面性，並聚焦香港K12機構面臨的具體威脅與解決方案。

K12教育中的 AI 網絡安全與數據安全簡介

數碼校園的雙重防線 - 安全與私隱

定義：

AI 網絡安全：保護系統、網絡、設備免受攻擊、破壞或未經授權存取。

AI 數據安全：確保學生、教職員敏感信息（如成績、身份資料）的保密性、完整性。

K12教育中的 AI 網絡安全與數據安全簡介

數碼校園的雙重防線 - 安全與私隱

香港教育現狀:

97%學校使用雲端教學平台（教育局數據）

BYOD政策普及→設備管理複雜化

K12教育中的 AI 網絡安全與數據安全簡介

數碼校園的雙重防線 - 安全與私隱

為何重要：

合規要求：違反《私隱條例》最高罰款100萬港元

信任基礎：家長對數據處理透明度的期待

教育領域的AI應用：機遇與風險

教育領域的AI應用：機遇與風險

AI革新教育—機會與暗礁並存

機遇

AI技術為教育帶來革命性突破：

個性化學習：

通過學習數據分析，AI可為學生提供適性化教材與即時反饋。

教學效率提升：

自動化評分、課堂管理工具減輕教師行政負擔。

教育普惠：

AI語言翻譯與輔助工具幫助少數族裔或特殊需求學生融入主流教育。

教育領域的AI應用：機遇與風險

AI革新教育—機會與暗礁並存

風險

數據隱私漏洞：

學生生物特徵（如人臉識別）、行為數據若遭洩露，可能被用於詐騙或定向攻擊。

算法偏見：

訓練數據不足可能導致AI推薦系統歧視特定學生群體。

系統濫用風險：

生成式AI（如ChatGPT）可能被學生用於作弊, 或產生虛假教育內容。

香港K12機構面臨的AI網絡安全主要威脅

香港K12機構面臨的主要 AI 威脅

香港學校的數碼 AI 防禦戰

威脅地圖

- AI 外部攻擊：勒索軟件（2023年教育界攻擊+40%）
- 內部疏失：教師,學生,家長誤開惡意郵件附件（佔事故65%）

香港K12機構面臨的主要 AI 威脅

香港學校的數碼 AI 防禦戰

AI 放大風險：

深度偽造語音冒充校長詐騙（2024年香港某國際學校案例）

自動化攻擊工具降低黑客門檻

May 8, 2025 10:00 PM Eastern Daylight Time

根據《2025年Imperva惡意機器人報告》，人工智慧推動了難以偵測的機器人的增加，目前這類機器人產生的流量已占全球網際網路流量的一半以上

Share

- 易於取得的AI工具的興起顯著降低了網路攻擊者的進入門檻，使他們能夠大規模建立和部署惡意機器人
- 自動化流量十年來首次超過了人類活動產生的流量，占所有網路流量的51%
- 針對應用程式介面(API)的攻擊在進階機器人流量中激增至44%，旅遊業整體上成為機器人攻擊的頭號目標

AI模型逆向工程攻擊

黑客試圖竊取訓練數據或模型參數，尤其威脅使用本地化粵語數據集的AI教育設施。

邊緣設備漏洞

智慧教室的IoT設備（如AI監控鏡頭、電子白板）常因未更新韌體成為入侵突破口。

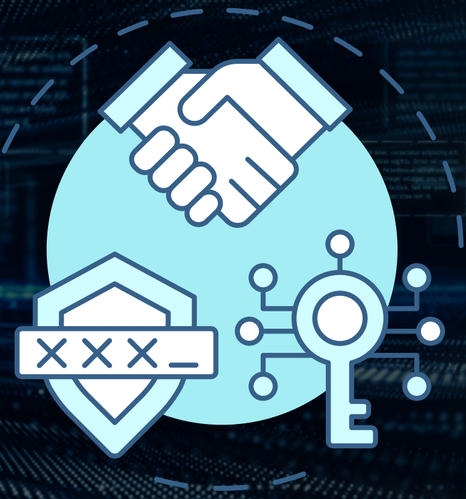
香港K12機構面臨的主要 AI 威脅

香港學校的數碼 AI 防禦戰

合規壓力：

- 跨境數據流動限制（如內地,外國《數據安全法》影響跨境平台）

跨境數據合規風險
使用境外AI服務（如中國內地或美國雲端平臺）
可能違反香港《個人資料（私隱）條例》。



跨境數據合規的本質是「法律管轄權競合」問題。香港K12機構應建立三層防護：

- 法律盡調：在使用內地AI服務前，要求供應商提供數據出境合規證明（如安全評估備案回執）；
- 技術隔離：通過數據脫敏、邊緣計算等方式減少原始數據出境；
- 透明度管理：向家長清晰說明數據跨境場景與權益保障措施，避免信任危機。

保障AI系統網絡安全的最佳實踐

保障AI系統網絡安全的最佳實踐

從防禦到韌性 - AI時代的安全架構

技術層面：

數據最小化：僅收集必要資料（如匿名化課堂參與度數據）

加密雙保險：傳輸加密(TLS 1.3)+靜態加密(AES-256)

零信任架構 (Zero Trust)： 將傳統資安實踐與AI專屬防護結合有效應對新型威脅

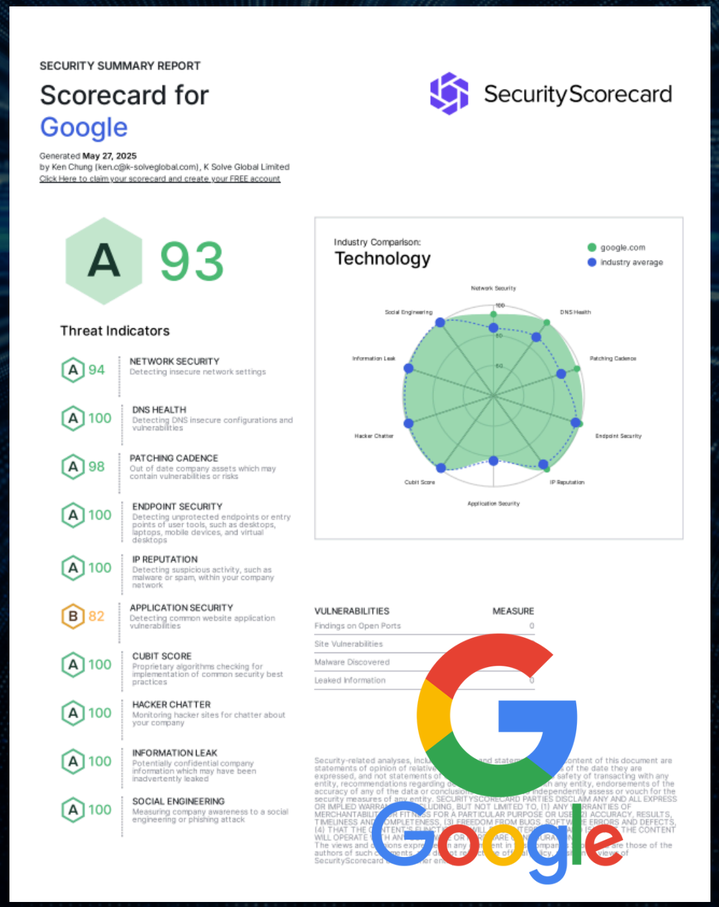
能在保護敏感資料、模型和基礎架構的同時，推動創新發展

保障AI系統網絡安全的最佳實踐

從防禦到韌性 - AI時代的安全架構

管理層面：
第三方風險評估表

[供應商名稱]
數據存儲地
ISO認證
年度滲透測試結果


SECURITY SUMMARY REPORT

Scorecard for
Google

SecurityScorecard

Generated May 27, 2025
by Ken Chung (ken.c@k-solveglobal.com), K Solve Global Limited
Click Here to claim your scorecard and create your FREE account

A 93

Industry Comparison:
Technology
google.com
industry average

Threat Indicators

A 94	NETWORK SECURITY	Detecting insecure network settings
A 100	DNS HEALTH	Detecting DNS insecure configurations and vulnerabilities
A 98	PATCHING CADENCE	Out of date company assets which may contain vulnerabilities or risks
A 100	ENDPOINT SECURITY	Detecting unprotected endpoints or entry points of user tools, such as desktops, laptops, mobile devices, and virtual desktops
A 100	IP REPUTATION	Detecting suspicious activity, such as malware or spam, within your company network
B 82	APPLICATION SECURITY	Detecting common website application vulnerabilities
A 100	CUBIT SCORE	Proprietary algorithms checking for implementation of common security best practices
A 100	HACKER CHATTER	Monitoring hacker sites for chatter about your company
A 100	INFORMATION LEAK	Potentially confidential company information which may have been inadvertently leaked
A 100	SOCIAL ENGINEERING	Measuring company awareness to a social engineering or phishing attack

VULNERABILITIES	MEASURE
Findings on Open Ports	0
Site Vulnerabilities	
Malware Discovered	
Leaked Information	0


SECURITY SUMMARY REPORT

Scorecard for
Deepseek

SecurityScorecard

Generated May 27, 2025
by Ken Chung (ken.c@k-solveglobal.com), K Solve Global Limited
Click Here to claim your scorecard and create your FREE account

A 93

Industry Comparison:
Unknown
deepseek.com
industry average

Threat Indicators

A 95	NETWORK SECURITY	Detecting insecure network settings
A 100	DNS HEALTH	Detecting DNS insecure configurations and vulnerabilities
A 100	PATCHING CADENCE	Out of date company assets which may contain vulnerabilities or risks
A 100	ENDPOINT SECURITY	Detecting unprotected endpoints or entry points of user tools, such as desktops, laptops, mobile devices, and virtual desktops
A 100	IP REPUTATION	Detecting suspicious activity, such as malware or spam, within your company network
B 88	APPLICATION SECURITY	Detecting common website application vulnerabilities
A 100	CUBIT SCORE	Proprietary algorithms checking for implementation of common security best practices
A 100	HACKER CHATTER	Monitoring hacker sites for chatter about your company
A 100	INFORMATION LEAK	Potentially confidential company information which may have been inadvertently leaked
A 100	SOCIAL ENGINEERING	Measuring company awareness to a social engineering or phishing attack

VULNERABILITIES	MEASURE
Findings on Open Ports	0
Site Vulnerabilities	16
Malware Discovered	0
Leaked Information	

2025 協作建立
第三方風險評估app

Grade	Company Name	Supplier CSRA	Certification
D	ESET Hong Kong	2025 | 2024 | 2023	ISO27001
C	FUJIFILM Business Innovation Hong Kong Limited	2025 | 2024 | 2023	ISO27001
B	Promethean Hong Kong	2025 | 2024 | 2023	ISO27001
B	Splashtop Hong Kong	2025 | 2024 | 2023	ISO27001
C	Blackboard Hong Kong	2025 | 2024 | 2023	ISO27001
B	NEC Hong Kong	2025 | 2024 | 2023	ISO27001
F	Konica Minolta Hong Kong	2025 | 2024 | 2023	ISO27001
A	Kodable Hong Kong	2025 | 2024 | 2023	ISO27001
B	Seesaw for Schools	2025 | 2024 | 2023	ISO27001
F	eClass (Modern Education Research Society)	2025 | 2024 | 2023	ISO27001
D	Tencent Education	2025 | 2024 | 2023	ISO27001
B	Canon Hong Kong	2025 | 2024 | 2023	ISO27001
B	Ricoh Hong Kong	2025 | 2024 | 2023	ISO27001
B	Lenovo Hong Kong	2025 | 2024 | 2023	ISO27001
C	HKBN (Hong Kong Broadband Network)	2025 | 2024 | 2023	ISO27001
F	iASPEC Education	2025 | 2024 | 2023	ISO27001
B	VTech Education	2025 | 2024 | 2023	ISO27001

Total Company	58
Grade A	13
Grade B	23
Grade C	11
Grade D	6
Grade F	5

保障AI系統網絡安全的最佳實踐

從防禦到韌性 - AI時代的安全架構

持續改進：

每學期「安全壓力測試」：模擬AI系統遭入侵場景

AI系統入侵應急響應流程報告

案例研究與經驗總結

案例研究與經驗總結

全球教育AI安全實戰

學生學習平台

分析學生作業

自動化行政任務

智能溝通工具

虛擬助教

沉浸式學習

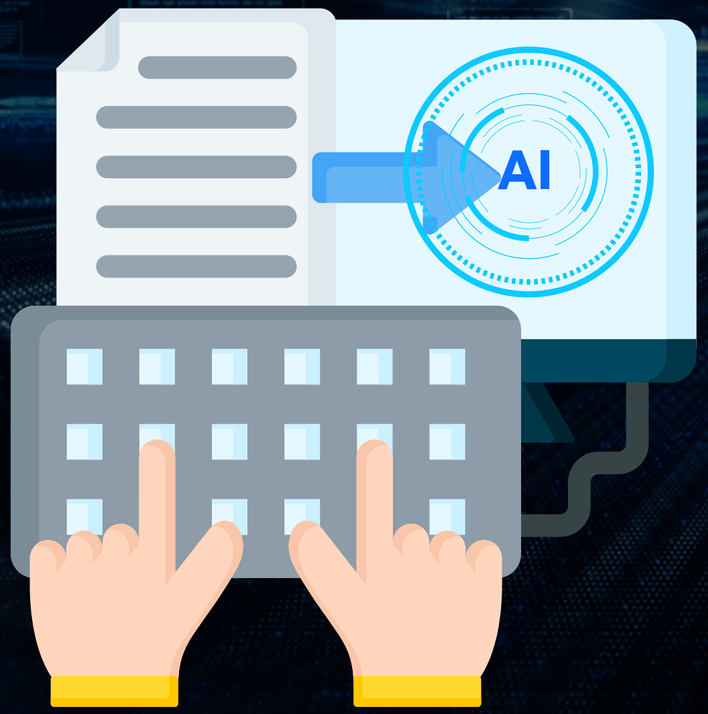
學習習慣分析

教學資源生成

教學分析與改進

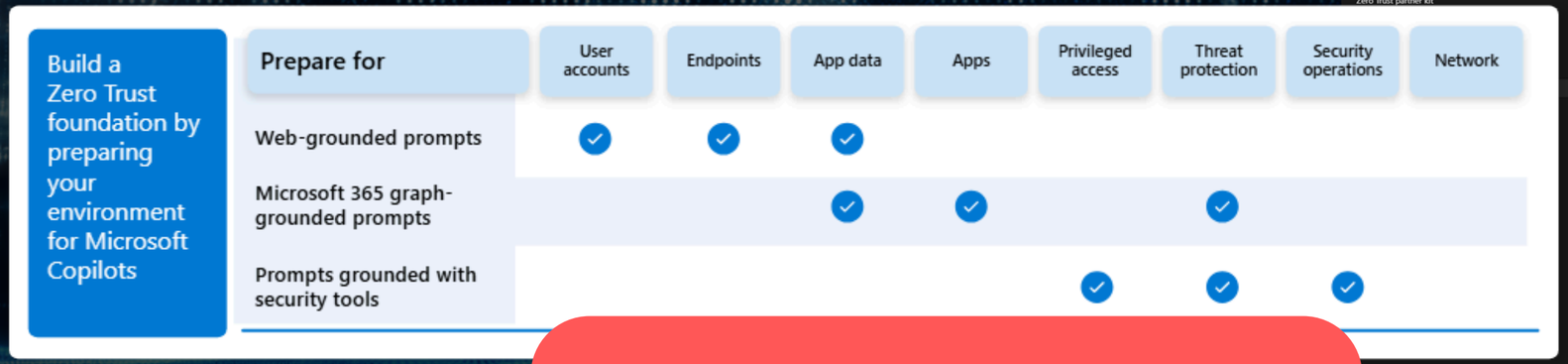
案例研究與經驗總結

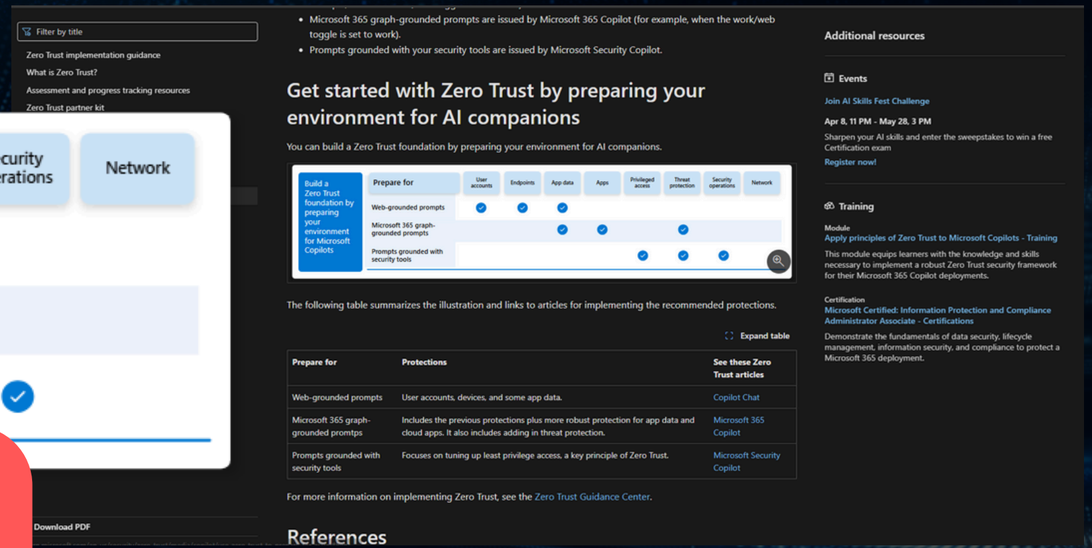
用之前想一想



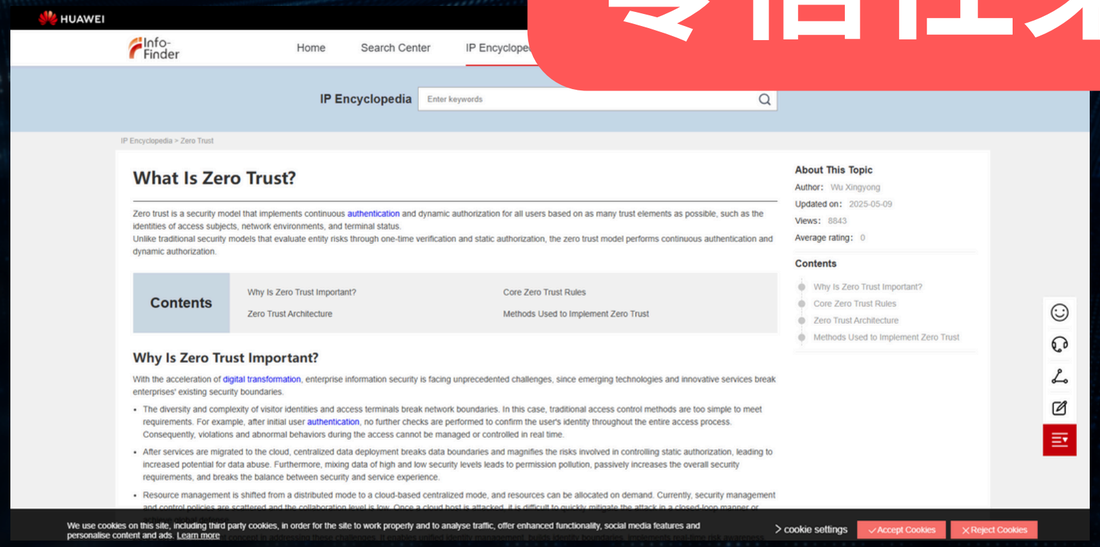
1. 學生個人身份資訊
2. 學業成績與評估紀錄
3. 健康與特殊需求資料
4. 家庭背景資訊
5. 行為紀律與輔導紀錄
6. 生物辨識資料
7. 地理位置與行蹤數據
8. 線上學習平台活動紀錄
9. 教職員人事資料

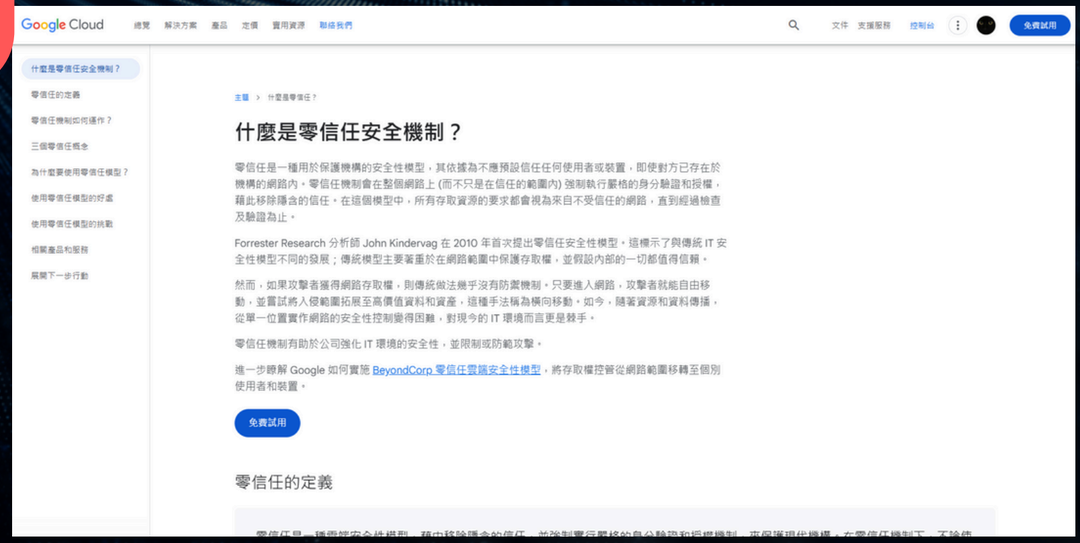
案例研究與經驗總結





零信任架構





案例研究與經驗總結

Poe

姓名	数学	语文	英语	总成绩	平均成绩
林明明	86	94	83	263	87.66666667
王进	79	78	79	236	78.66666667
杨柳书	94	81	84	259	86.33333333
刘小诗	82	98	91	271	90.33333333
刘诗琦	78	82	94	254	84.66666667
元星	80	86	75	241	80.33333333
李光	69	79	85	233	77.66666667
郭敏浩	58	83	76	217	72.33333333

DeepSeek-R1-FW
由 @fireworksai

可信任平台?

案例研究與經驗總結

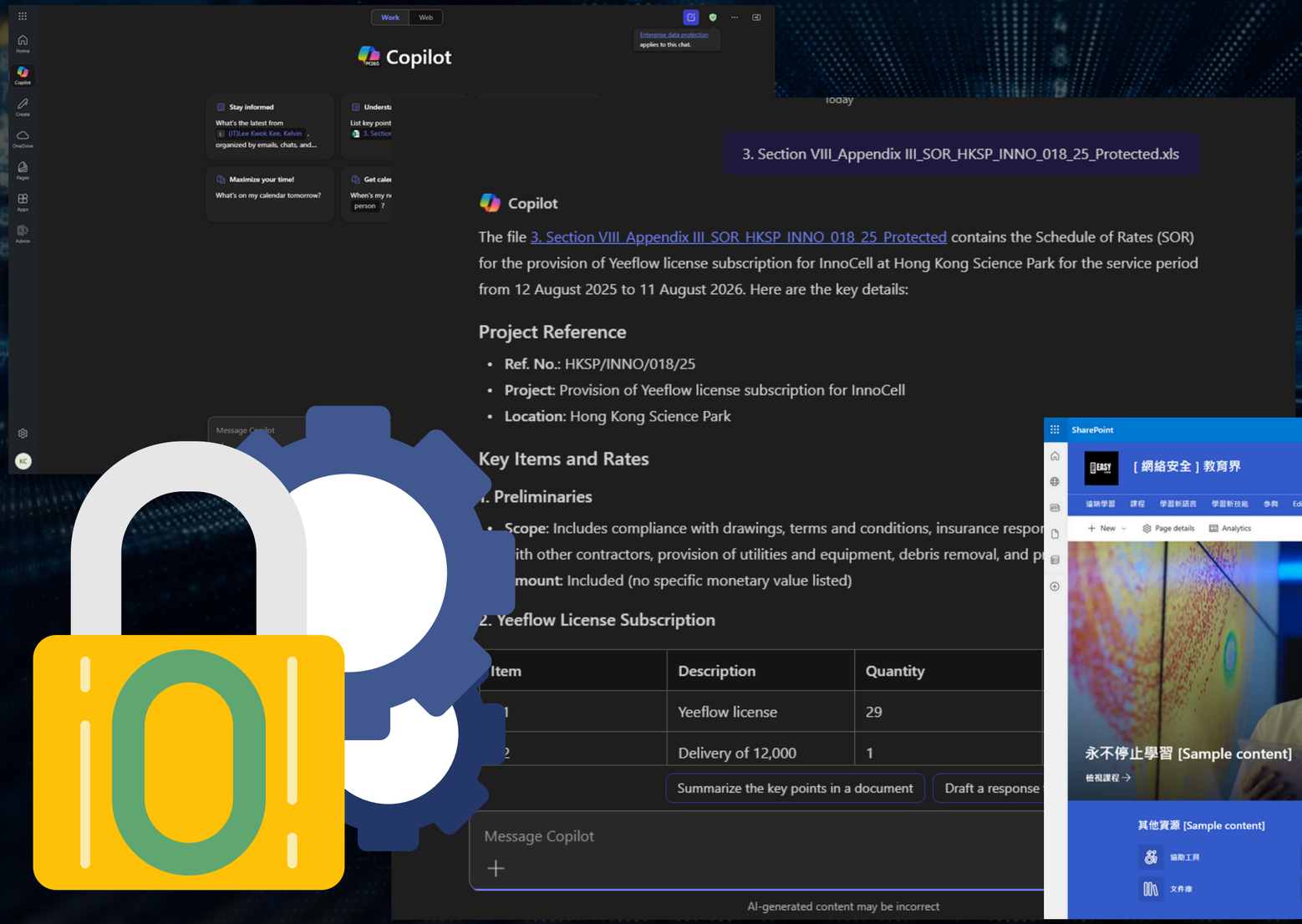
Poe

姓名	数学	语文	英语	总成绩	平均成绩
林明明	86	94	83	263	87.66666667
王进	79	78	79	236	78.66666667
杨柳书	94	81	84	259	86.33333333
刘小诗	82	98	91	271	90.33333333
刘诗琦	78	82	94	254	84.66666667
元星	80	86	75	241	80.33333333
李光	69	79	85	233	77.66666667
郭敏浩	58	83	76	217	72.33333333

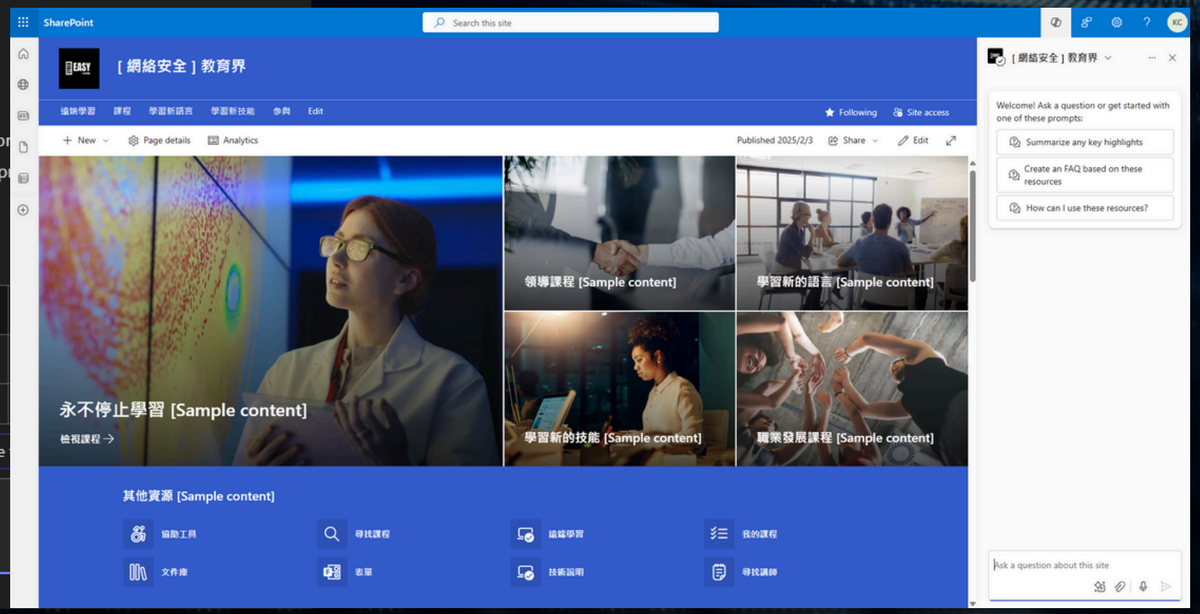
處理敏感數據

姓名	数学	语文	英语	总成绩	平均成绩
001	86	94	83	263	87.66666667
002	79	78	79	236	78.66666667
003	94	81	84	259	86.33333333
004	82	98	91	271	90.33333333
005	78	82	94	254	84.66666667
006	80	86	75	241	80.33333333
007	69	79	85	233	77.66666667
008	58	83	76	217	72.33333333

案例研究與經驗總結


Work
Web
Copilot
Stay informed
What's the latest from
organized by emails, chats, and...
Maximize your time!
What's on my calendar tomorrow?
Enterprise data protection
applies to this chat.
3. Section VIII_Appendix III_SOR_HKSP_INNO_018_25_Protected.xls
Copilot
The file 3. Section VIII_Appendix III_SOR_HKSP_INNO_018_25_Protected contains the Schedule of Rates (SOR) for the provision of Yeeflow license subscription for InnoCell at Hong Kong Science Park for the service period from 12 August 2025 to 11 August 2026. Here are the key details:
Project Reference
Ref. No.: HKSP/INNO/018/25
Project: Provision of Yeeflow license subscription for InnoCell
Location: Hong Kong Science Park
Key Items and Rates
1. Preliminaries
2. Yeeflow License Subscription
Item
Description
Quantity
Yeeflow license
29
Delivery of 12,000
1
Summarize the key points in a document
Draft a response
Message Copilot
AI-generated content may be incorrect


Enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat
Article • 03/13/2025 • 2 contributors
Feedback
In this article
What is enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat?
Enterprise data protection for prompts and responses
Privacy and security of web queries
Agents in Microsoft 365 Copilot
Footnotes
What is enterprise data protection in Microsoft 365 Copilot and Microsoft 365 Copilot Chat?
The use of Microsoft 365 Copilot and Microsoft 365 Copilot Chat, as used by organizations, are covered by the terms of the Data Protection Addendum (DPA) and Product Terms, with Microsoft acting as a data processor.
Enterprise data protection (EDP) refers to controls[1] and commitments, under the Data Protection Addendum (DPA) and Product Terms, that apply to customer data for users of Microsoft 365 Copilot and Microsoft 365 Copilot Chat. The use of the term EDP isn't meant to limit the benefits offered under the DPA and Product Terms.


SharePoint
[網絡安全] 教育界
永不停止學習 [Sample content]
領導課程 [Sample content]
學習新的語言 [Sample content]
學習新的技能 [Sample content]
職業發展課程 [Sample content]
其他資源 [Sample content]
Welcome! Ask a question or get started with one of these prompts:
Summarize any key highlights
Create an FAQ based on these resources
How can I use these resources?
Ask a question about this site

護數守設
安全之道

歡迎聯絡我