

生成AIの誤判断とデジタル社会 の生命線と企業・国家リスク

人工知能が、データを検閲している。日本のクラウド利用のリスク議論する

日本インターネットガバナンスフォーラム (JapanIGF) 2024、2024年11月5日 (火)

中澤祐樹

一般社団法人ネット情報信頼性機構 理事

ライトセンド株式会社代表 (ICANN公認レジストラ、ウルトラドメイン運営)

スカイクリア株式会社取締役 (ICANN会員、バックオーダー事業者)

中澤祐樹

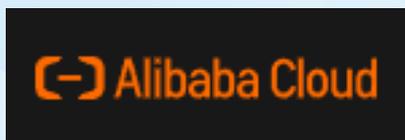


一般社団法人ネット情報信頼性機構

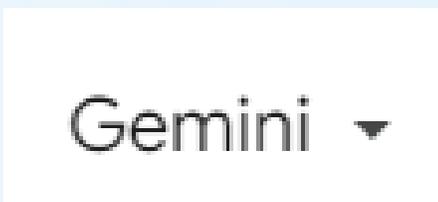
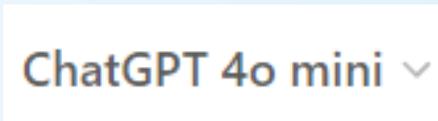


身近になったクラウドサービスと生成AI

クラウドサービス



生成AI・人工知能



身近になったクラウドサービスと生成AIと問題

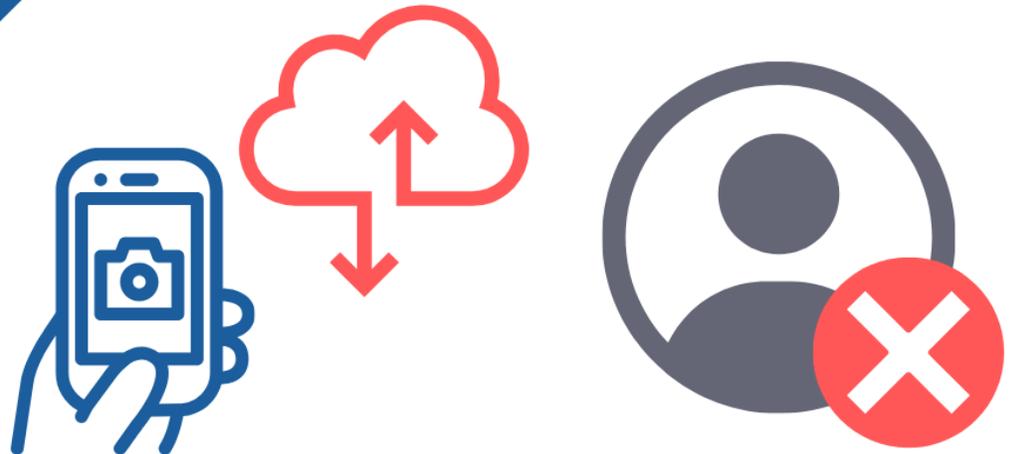
Combat fraud using AI

オンライン詐欺と戦うAI技術

という表題のスライド資料を
スマートフォンで撮影



撮影したスライド写真を
OneDriveに保存したところ
アカウントが凍結される…



注意を要した方が良い写真



AI 科技打詐 Combat fraud using AI

- Chungwa Telecom, as the leading telecom brand in Taiwan, responds to the Executive Yuan's "New Generation Anti-Fraud Strategy Action Plan 1.5" by establishing the Anti-Fraud PMO team in 2023. The team **actively blocks fraud at the source and continues to lead in launching technological anti-fraud measures**. These measures include blocking overseas fraudulent voice calls and domestic and international fraudulent SMS messages, providing interception and warnings from fixed networks to mobile networks, and continuously expanding the scope of protection.
- Through AI technology to combat fraud, according to the "2023 Annual Report" released by the CIB (Criminal Investigation Bureau) in collaboration with Whoscall, Taiwan's anti-fraud efforts in 2023 have been successful. The interception of overseas fraudulent calls and voice warnings have been effective, **making Taiwan the country with the most significant decrease in fraudulent calls and SMS messages in Asia**.



- 誠信為本
- 客戶信賴
- 創新創價
- 承諾當先

- 中華電信身為台灣電信領導品牌，響應行政院「新世代打擊詐欺策略行動綱領 1.5」，2023年率先成立打詐PMO小組，**積極源頭堵詐，持續領先推出科技防詐**，防堵境外詐騙語音及境內外詐騙簡訊，從攔截到警示、從固網到行網，持續擴大防護面。
- 透過AI科技打詐，根據刑事局偕同Whoscall公布的「2023年度報告」，2023年台灣打詐有成，境外詐騙電話攔阻及語音警示有成效，是亞洲各國詐騙電話及簡訊下降最多的國家。

注意を要した方が良い写真の会場風景 1



注意を要した方が良い写真の会場風景 2



マイクロソフトサービス規約違反検知により アカウント停止



nkzw

ご使用のアカウントがロックされました

Microsoft サービス規約に違反するアクティビティが検出されたため、アカウントをロックしました。

アカウントのロック解除

アカウントのロックを解除するには追加のサポートが必要です。["https://aka.ms/compliancelock"](https://aka.ms/compliancelock) に移動すると、適切な場所へ移動します。

アカウントがロックされている場合



The screenshot shows the Microsoft support website with the title "アカウントがロックされている場合" (Account locked). The navigation bar includes "Microsoft", "サポート" (Support), "Microsoft 365", "Office", "製品" (Products), "デバイス" (Devices), "アカウントと請求" (Accounts and Billing), "リソース" (Resources), and a button "Microsoft 365 を購入" (Buy Microsoft 365). The main heading is "アカウントがロックされている場合" (Account locked) with the sub-heading "Microsoft account". The text explains that an account is locked if it has violated the terms of service. It provides instructions on how to verify the account status using an online form and mentions that Microsoft will send a ticket number via email and provide further assistance through email or a request for more details. A link for the "Microsoft locked account review form" is provided at the bottom of the text area.

解除申請フォーム（英語のみ）

 Take the power of AI on the go with the free Copilot app
Create images, get help with writing, and search faster

[No, thanks](#) [Get the Copilot app](#)

 Microsoft

Microsoft 365 Teams Copilot Windows Surface Xbox Deals Small Business Support

All Microsoft Search Cart

Complete the form below for Microsoft to review the reason your account was disabled?

There are multiple reasons your account may have been disabled including suspicious activity, sending unsolicited emails, or for violating the [violationTypeLink](#) by hosting photos, video or other content in violation of the [Code of Conduct](#). To request that we review the reason your account was disabled and determine whether it may be reinstated, complete and submit the form below.

* Indicates required fields

Your name*

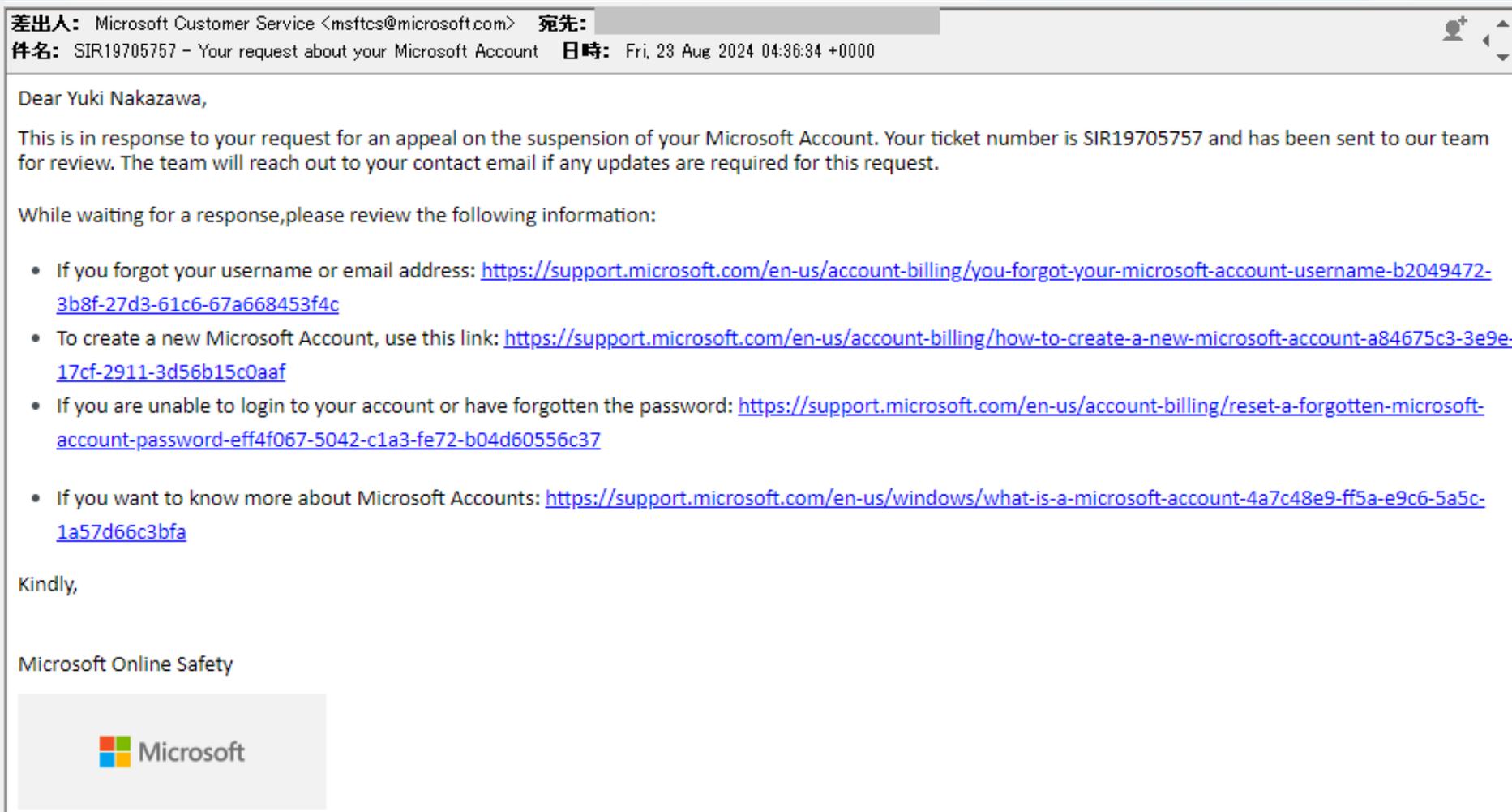
Contact email address*

Microsoft may not need to contact you. Microsoft will determine the validity of your request, as well as any follow-up action, but may not contact you with updates about your request or any action. Note that if Microsoft detects an email address is being used to abuse this reporting process for the purpose of harming a Microsoft customer or Microsoft services, or for other improper purposes, we may block the email address from submitting reports.

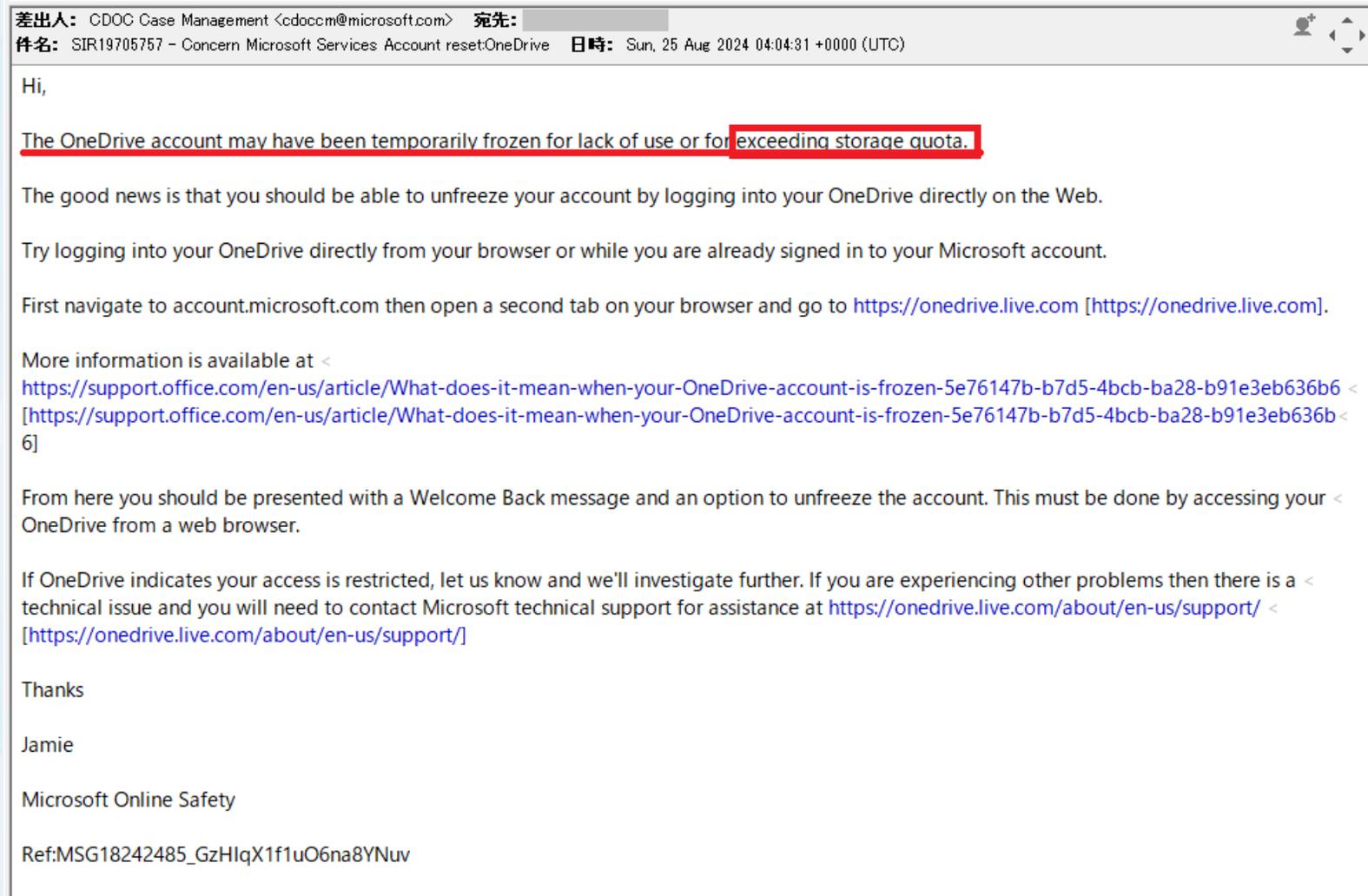
Disabled account email address*

The primary email address, member ID, or Skype username associated with the disabled account to be reviewed.

解除申請受領メール



ディスク容量超過のためアカウント停止と返信が来る



ディスク容量超過の事実はない

The screenshot shows the '自分' (Me) storage settings page on an iPhone. At the top, the status bar displays the time 13:26, location, and battery level at 99%. Below the header, the user's name 'nkzw' is shown. A red box highlights the storage usage: '1 TB (89%) のうち 938.6 GB を使用'. Below this is a progress bar showing approximately 89% usage. A list of options is visible at the bottom: 'オフラインで使用可能なファイル', 'ごみ箱', '通知', '設定', and 'ヘルプとフィードバック'.

今回の要点

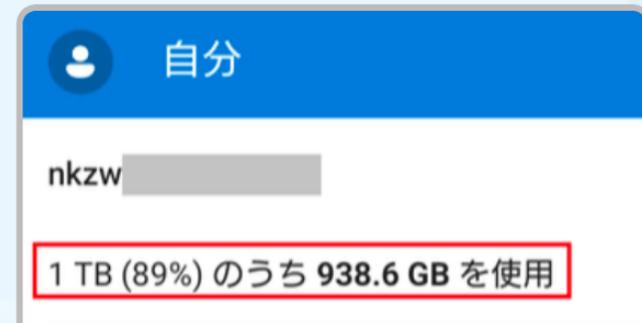
ユーザー数を考えると
機械判定によるものだと
考えられる…？



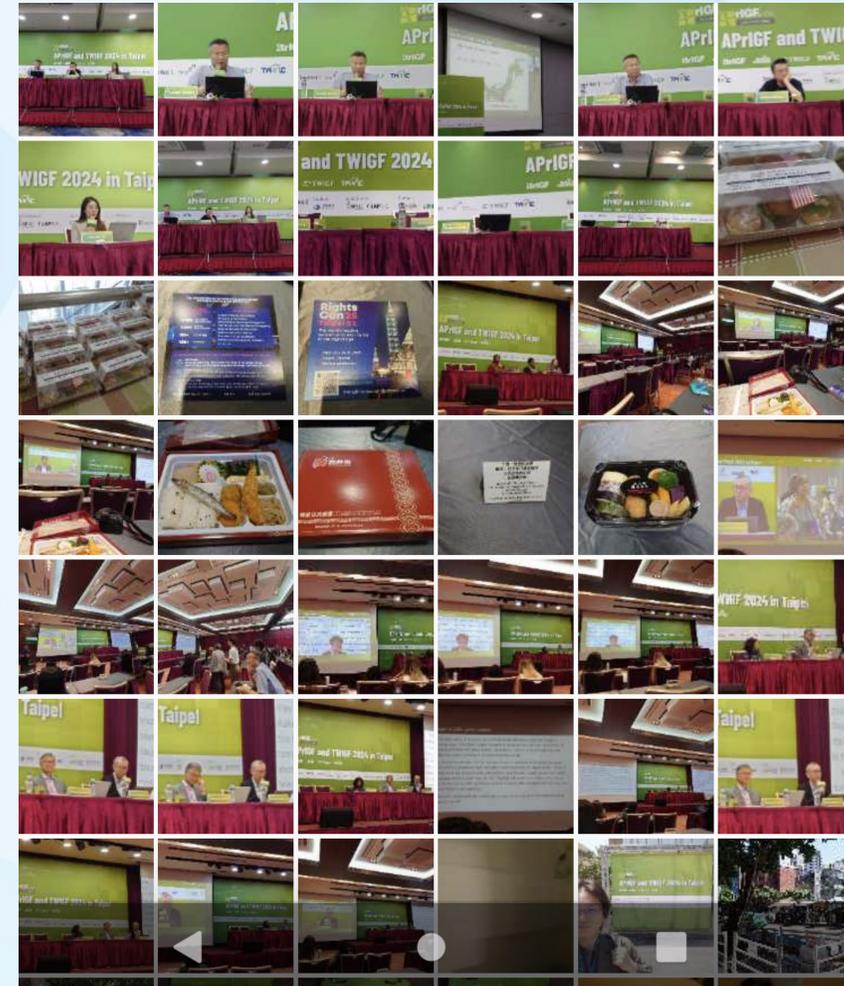
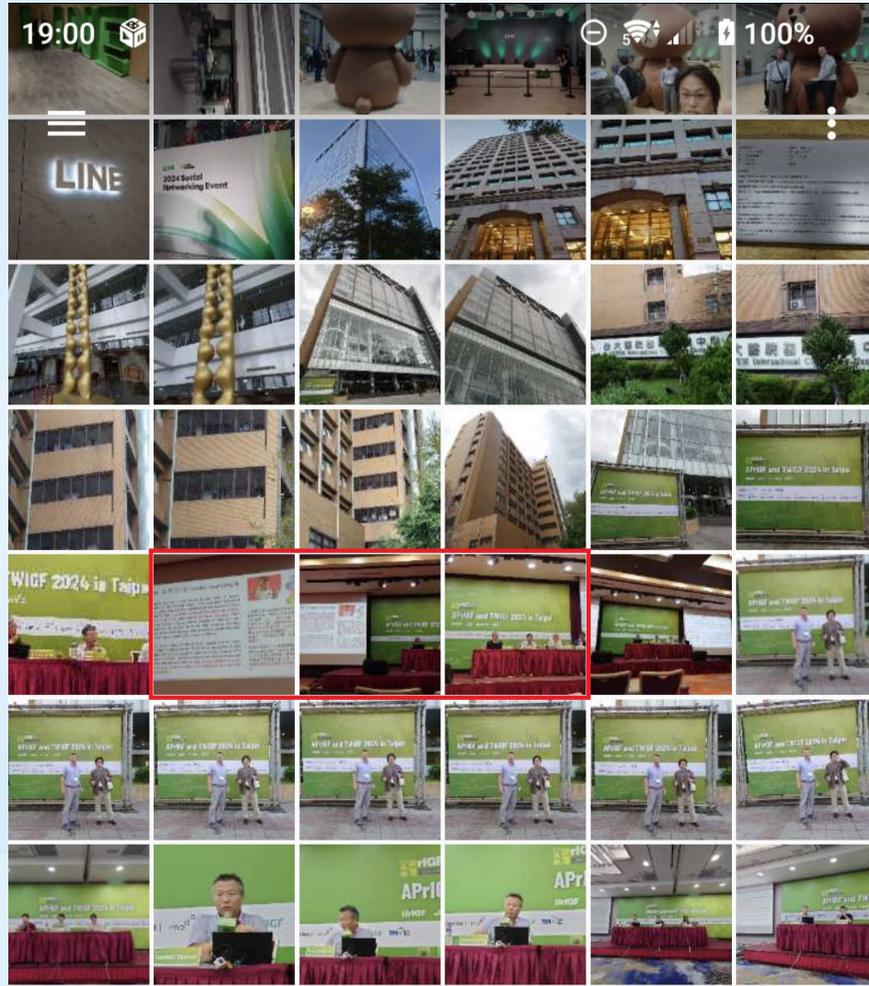
AIによってfraud(詐欺)が
引っかけた可能性…？
この時点で事業者に検閲されている



問い合わせすると容量超過とのこと…
実際は容量超過していない状態

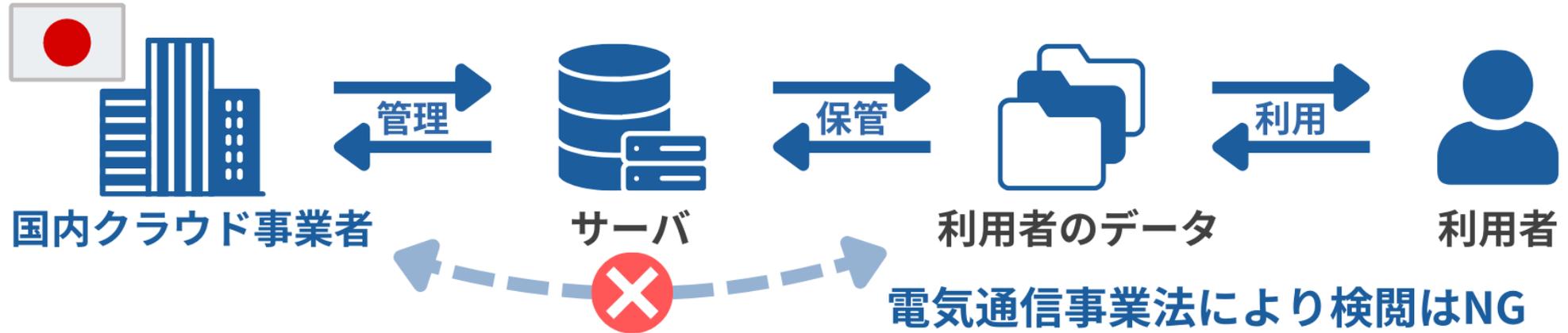


2024年8月22日（木）に撮影した写真一覧



生成AIの検閲と通信の秘密が守られていない

国内クラウド事業者はデータの検閲に電気通信事業法の制限がある



海外クラウド事業者にはデータ検閲の制限が適用されていない懸念



国家機密や企業の機密情報の漏洩の可能性

海外クラウドサービスを利用している場合
知らずにデータを検閲されてしまう可能性が高い

大手企業



取引文書
契約書
研究情報など…



政府/国家



軍事資料
外交文書
国家機密など…



海外クラウド事業者にて
データを閲覧される可能性



海外クラウド事業者にて
管理されているクラウドサーバ

日本国内の重要な情報が海外へ流出してしまうことで
外交上の不利/国家間の圧力へ繋がってしまう場合も想定される…