Cloud

KI und das Ende der pseudonymen Online-Identität

Die Möglichkeit, Menschen anhand vermeintlich anonymer oder pseudonymer Daten zu identifizieren, ist nicht neu. Doch erschreckend ist, wie schn einfach und zugänglich die Deanonymisierung mit Hilfe der KI geworden ist.

By: Udo Schneider
Mar 04, 2026
Read time: 4 min (978 words)

„Even while we sleep / We will find you." Wer kennt diesen Ohrwurm aus „Everybody Wants to Rule the World" von Tea Fears nicht? Damals, 1985, war das eine Metapher für den Kalten Krieg und das Machtstreben. Heute, im Februar 202 ich diese Zeile und denke: Das ist keine Metapher mehr. Es ist eine technische Beschreibung. Angreifer können heute tatsächlich schlafen, weil KI die Aufklärungsarbeit automatisch und rund um die Uhr erledigt.

Das Team von TrendAI™ Research hat kürzlich demonstriert, wie einfach es ist, aus öffentlichen LinkedIn-Daten vollautomatisch eine maßgeschneiderte Phishing-Kampagne zu erstellen: ein PoC-Tool, gebaut von einem einzigen Entwickler in unter 30 Minuten, liefert personalisierte Phishing-Seiten, die auf die gesamte Führungsebene eines Unternehmens zielen. Erwartungsgemäß gab es in den Kommentaren die Reaktion: „LinkedIn ist mein professionelles und das sagt ja nichts über mich als reale Person aus." Diese Einschätzung ist verständlich, aber leider falsch!

Die Illusion der getrennten Identitäten

Die unbequeme Wahrheit: Die Möglichkeit, Menschen anhand vermeintlich anonymer oder pseudonymer Daten zu identifizieren, ist nicht neu. Die US-amerikanische Informatikerin Latanya Sweeney hatte bereits um die Jahrtausendwe mit Hilfe der US-Volkszählungsdaten gezeigt, dass 87% der US-Bevölkerung allein anhand von Postleitzahl, Geburtsdat und Geschlecht eindeutig identifiziert werden können.

In einer neueren Studie von Swanson, Lermen et al. (ETH Zürich, Anthropic) machen die Forscher deutlich, dass LLMs anonyme Online-Profile mit erschreckender Präzision plattformübergreifend mit realen Identitäten verknüpfen könne einem ihrer Experimente wurden 67 % von 338 Zielpersonen korrekt aus Hacker-News-Aktivitäten auf ihr Linke Pro zurückgeführt, vollautomatisch, bei 90 % Präzision und 1 bis 4 US-Dollar pro Zielperson. Die sorgfältig getrennten Onli Identitäten existieren also faktisch nicht mehr, zumindest nicht aus der Perspektive eines entschlossenen Angreifers.

Die Studie der ETH Zürich und Anthropic zeigt, dass LLMs nicht auf strukturierte Daten oder vordefinierte Merkmalsschemata angewiesen sind. Sie extrahieren Identitätssignale direkt aus unstrukturiertem Text: Schreibstil, beiläufige Ortsangaben, Interessen, demografische Angaben, akademischer Hintergrund… Alles, was ein Mensch aus e Gespräch mit jemandem lernen würde, lernt das Sprachmodell aus Posts und Kommentaren.

Was sich mittlerweile grundlegend geändert hat, ist nicht *was* möglich ist, sondern wie *schnell* und mit *welchem Aufwar* heute machbar ist!

Wenige Datenpunkte, große Wirkung

Es lohnt sich, über ein klassisches Phänomen aus der Wahrscheinlichkeitsrechnung (das Geburtstagsparadoxon) nachzudenken: Wie viele Personen müssen sich in einem Raum befinden, damit die Wahrscheinlichkeit größer als 50 ' dass *zwei beliebige* Personen am selben Tag Geburtstag haben? Die intuitive Antwort liegt bei 100 oder mehr. Die mathematisch korrekte Antwort: *23 Personen*. Bei 50 Personen liegt die Wahrscheinlichkeit bereits über 97 %.

Das zeigt, dass Menschen systematisch unterschätzen, wie schnell kombinatorische Wahrscheinlichkeiten ansteigen: M jedem weiteren öffentlichen Datenpunkt wächst die Zahl möglicher Verknüpfungen kombinatorisch, nicht linear.

Übertragen auf die Online-Deanonymisierung ergibt sich folgendes Bild: Jeder einzelne Datenpunkt erscheint harmlos LinkedIn-Post zu einer Konferenz, ein Urlaubsfoto auf Instagram, ein Reddit-Kommentar zu einem Film. Keiner dieser Datenpunkte kann für sich genommen eine Person eindeutig identifizieren.

Aber kann *irgendjemand* aus einer Datenbank von Millionen Profilen mit *irgendeiner* Kombination von Datenpunkten identifiziert werden? Diese Wahrscheinlichkeit ist erschreckend hoch. Je mehr Daten zur Verfügung stehen, desto einfa wird die Deanonymisierung.

Dreißig öffentliche Instagram-Bilder reichen aus, um in unter 30 Minuten eine maßgeschneiderte Phishing-Kampagne erstellen. Was früher eine Woche Arbeit zweier Forscher erforderte, erledigt heute das erwähnte PoC-Tool in einer hal Stunde. Standort, Gewohnheiten, Gesundheitszustand, politische Neigungen: alles extrahiert aus Metadaten und Bildi vollautomatisch. Ein harmloser Urlaubsschnappschuss ist aus Angreiferperspektive ein Datenpunkt.

Fazit

Dass Verhaltensdaten in Kombination ein präzises Identitätsprofil ergeben, ist seit 2016 öffentlich bekannt, spätestens Grassegger und Krogerus die Cambridge-Analytica-Geschichte für *Das Magazin* aufbrachen. Kosinski, dessen Forschur Cambridge Analytica zugrunde lag, brachte es auf den Punkt: *„Ich habe die Bombe nicht gebaut. Ich habe nur gezeigt, da sie gibt."* War die Bombe 2016 groß, schwer und erforderte eine gut ausgerüstete Organisation, um sie zu transportier ist sie nun zehn Jahre später handgepäcktauglich. Es geht darum, *wie schnell* und *mit wie wenig Aufwand* sie heute durchführbar sind.

Kombiniert heißt das, dass Angreifer Daten (inkl. Photos) aus verschiedenen Quellen, egal ob beruflich oder privat, dur LLMs automatisiert zusammenführen können und dann (ebenfalls automatisiert) absolut zielgerichtete Phishing-Kampagnen starten können … während sie schlafen …

Die wichtige Erkenntnis für Sicherheitsteams: **Ein Mitarbeiterdatenrisiko ist ein Unternehmensrisiko.** Und für dies geschlossenen Kreislauf haben die meisten Unternehmen noch kein Sicherheitskonzept, weil der Übergang zwischen privatem und beruflichem digitalem Fußabdruck bisher kein operatives Sicherheitsproblem war. Heute ist er es.

Wer Risikomodelle noch auf der Prämisse aufbaut, bestimmte Angriffe seien „zu aufwändig", sollte diese überdenken. limitierende Faktor ist heute nicht mehr die Expertise, sondern die Motivation.

Was also tun?

Keine Panik, aber Klarheit. Einige konkrete Überlegungen:

- Footprint-Bewusstsein schaffen: Mitarbeiter sollten verstehen, dass ihr öffentlich sichtbares Online-Verhalten Teil der Angriffsfläche ihres Arbeitgebers ist.
- Exposure Management überdenken: Unternehmen sollten nicht nur ihre IT-Infrastruktur, sondern auch den externen digitalen Fußabdruck Belegschaft in ihr Sicherheitskonzept integrieren.
- Bedrohungsmodell aktualisieren: Die Kostenannahme für personalisierte Angriffe ist weggefallen. Risikoabschätzungen, die auf manuellem Aufwand als Bremse beruhen, sind veraltet.
- Bildinformationen ernst nehmen: Fotos tragen mehr Informationen in sich als sichtbar. EXIF-Daten entfernen, Bildanalyse-Potenzial verste Das Urlaubsfoto ist kein privates Dokument mehr: Es ist ein (weiterer) Datenpunkt.
- Regulatorischen Rahmen kennen: Die DSGVO schränkt die automatisierte Profilbildung von natürlichen Personen unter bestimmten Bedingungen ein; der AI Act wird diesen Bereich weiter adressieren. Für Angreifer ist das irrelevant! Für Unternehmen als Arbeitgeber jedo wichtiger Orientierungspunkt beim Umgang mit Mitarbeiterdaten.
- Überprüfung zeitbasierter Schutzannahmen: Angriffe, die früher Wochen dauerten, dauern heute Minuten.

Tags

Expertenmeinung | Compliance und Risiko | Endpunkte | Ransomware | Ausnutzung von Schwachstellen | Social Media | Deep Web | ASRM | Risikomanagement | Rechenzentrum | Bericht | Cloud | Malware | Web | Phishing | Connected Car | Plattform | TM Vision One Platform | Cyberrisiken | Cyber-Kriminalität | Smart Home | Risiken für die Privatsphäre | Cyberbedrohungen | ICS OT | APT und gezielte Angriffe | IoT | Künstliche Intelligenz (KI) | Spam | Mobilgeräte | Network

Authors

Udo Schneider
IoT Security Evangelist Europe

CONTACT US

Related Articles

Bösartige OpenClaw Skill täuschen Nutzer

Europol, Microsoft, TrendAI™ and Collaborators Halt Tycoon 2FA Operations

KI für die rasche Erstellung von Zielprofilen für Cyberkriminelle

See all articles ›

Select a language

Deutsch

Datenschutz
Barrierefreiheit
Sitemap

Rechtliches
Nutzungsbedingungen