

AKTE 03-A / SI-OR-2022-001

SYNTHETISCHE IDENTITÄTEN

OPERATIVE RISIKEN

VERFASSERIN	M. VOSS
STAND	DEZEMBER 2022
FASSUNG	1
STATUS	ARBEITSSTAND
VERTEILER	BEGRENZT

INTERN // NICHT ZUR EXTERNEN WEITERGABE

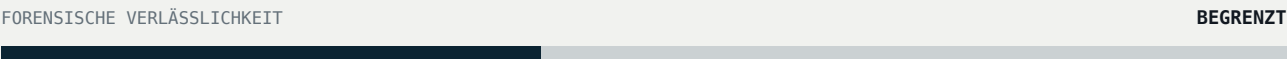
Zweck und Kurzbewertung

Dieses Arbeitspapier bündelt frühe operative Beobachtungen zu Identitäten, deren sichtbare Bestandteile ganz oder teilweise synthetisch erzeugt wurden. Es beschreibt kein neues Täterprofil. Es beschreibt eine veränderte Kostenstruktur für bekannte Delikte.

KERNBEFUND: Erzeugte Identitäten senken den Aufwand für glaubwürdige Täuschung und erhöhen gleichzeitig Reichweite, Anpassungsfähigkeit und Wiederverwendbarkeit.

BETRACHTETE KOMPONENTEN

KOMPONENTE	BEOBACHTUNG 2022	OPERATIVES RISIKO
Bild	Konsistente Personenbilder sind mit geringem Aufwand erzeugbar.	Identitäten können ohne reale Bildquelle aufgebaut werden.
Stimme	Transformation und Klonen sind fallweise bereits in Gesprächen nutzbar.	Emotionale Nähe und Autorität lassen sich glaubwürdig simulieren.
Text	Stil, Sprache und Rollenbezug sind automatisierbar.	Viele individualisierte Dialoge können parallel geführt werden.
Dokument / Video	Qualität schwankt; Einzelprüfungen bleiben fehleranfällig.	Gefälschte Belege stützen eine bereits etablierte Identität.
Verhalten	Profile können über Zeit konsistent gepflegt werden.	Vertrauen entsteht nicht nur durch Inhalte, sondern durch Wiederholung.



Bewertung basiert auf einem kleinen Fallbestand, Laborrückmeldungen und offenen technischen Beobachtungen. Die Richtung ist belastbarer als jede einzelne Prozentangabe.

01 / FALLVERBUND

Beobachtetes Muster

Ausgangspunkt ist ein Betrugsfall mit einer vollständig erfundenen Kontaktidentität. Das Opfer überwies innerhalb von acht Wochen rund 140.000 Euro auf mehrere Konten. Profilbilder, Lebensgeschichte und Kommunikation wirkten zusammenhängend; eine reale Person hinter der dargestellten Identität konnte nicht festgestellt werden.

REKONSTRUIERTER ABLAUF

PHASE	BEOBACHTUNG	RELEVANZ
01 Aufbau	Mehrere Bilder derselben vermeintlichen Person in Alltagssituationen.	Kleine Inkonsistenzen erst im Direktvergleich sichtbar.
02 Bindung	Wochenlange persönliche Textkommunikation, abgestimmt auf Reaktionen des Opfers.	Hoher Zeitaufwand kann teilweise automatisiert worden sein.
03 Stimme	Zwei kurze Telefonate ohne Video; spontane Reaktion auf eine Bemerkung.	Echtzeitanteil oder menschlicher Operator nicht trennscharf bestimmbar.
04 Transfer	Zahlungen an mehrere Identitäten und Konten in drei Ländern.	Klassische Verschleierung profitiert von digitaler Skalierung.
05 Wiederverwendung	Ähnliches Stimmuster in einem zweiten Verfahren mit anderer Rolle und Sprache.	Hinweis auf denselben Generierungsprozess, nicht zwingend denselben Sprecher.

DETEKTORERGEBNISSE

Drei Bilddetektoren bewerteten dasselbe Material mit 90 Prozent, 58 Prozent und 'wahrscheinlich echt'. Ein externes Labor sah bei zwei Audiodateien eine hohe Wahrscheinlichkeit synthetischer Erzeugung und möglicherweise denselben Generierungsprozess.

OPERATIVE FOLGE: Die Prüfung sichtbarer Fehler reicht nicht aus. Wo Herkunft entscheidend ist, muss Herkunft nachgewiesen werden.

OFFENE FRAGEN

- Welcher Anteil der Kommunikation wurde live von einem Menschen gesteuert?
- Welche Bestandteile wurden einmal erzeugt und anschließend wiederverwendet?
- Wie viele parallele Zielpersonen wurden durch dieselbe Infrastruktur angesprochen?
- Welche Signale bleiben stabil, wenn Bild, Stimme und Text gemeinsam variieren?

02 / MODELL

Technische Angriffsarchitektur

Synthetische Identität ist kein einzelnes Medium. Sie entsteht als Kette aus Datensammlung, Darstellung, Interaktion und Wiederholung. Jede Stufe kann menschlich, automatisiert oder gemischt betrieben werden.

KETTE

01	DATEN	Öffentliche Profile, Kontakte, Routinen, Sprache, aktuelle Ereignisse
02	AUFBAU	Name, Biografie, Bilder, Stimme, Dokumente, Konten und Geräte
03	AUSWAHL	Zielgruppen, Verwundbarkeit, Rolle, Zeitpunkt und geeigneter Kanal
04	INTERAKTION	Personalisierte Texte, Rückfragen, Erinnerungen, emotionale Reaktion
05	AUSLÖSER	Zahlung, Zugang, Information, Handlung oder öffentliche Weitergabe
06	SKALIERUNG	Parallele Gespräche, Übersetzung, Wiederverwendung und Optimierung

WAHRSCHEINLICHE BETRIEBSFORMEN

FORM	MENSCHLICHER ANTEIL	SYSTEMANTEIL	EINSCHÄTZUNG
Assistiert	Zielwahl, Freigabe, schwierige Dialoge	Entwürfe, Bilder, Übersetzung	bereits verbreitet
Überwacht	Ausnahmen, Zahlungsphase, Eskalation	laufende Dialoge, Profilpflege	kurzfristig plausibel
Weitgehend autonom	Zielsetzung, Grenzen, Verwertung	Auswahl, Kontakt, Anpassung, Übergabe	nicht belegt; vorbereitet

ENTSCHEIDENDER MULTIPLIKATOR

Ein einzelner Angreifer muss nicht mehr Bilder erzeugen, Texte formulieren, Sprachen beherrschen und jeden Dialog persönlich führen. Er muss zunehmend nur noch Ziel, Grenzen und gewünschte Handlung festlegen.

Der Engpass verschiebt sich von der Herstellung glaubwürdiger Kommunikation zur Auswahl lohnender Ziele und zur sicheren Verwertung des Ergebnisses.

03 / LAGEBILD

Risikoszenarien 2023-2025

Die folgenden Szenarien sind keine Prognosen über einzelne Werkzeuge. Sie beschreiben, welche bekannten Angriffe durch sinkende Herstellungs- und Betreuungskosten wahrscheinlicher werden.

SZENARIO	MECHANISMUS	FOLGE	RISIKO
Finanzbetrug	Persistente Beziehung mit Bild, Text und Stimme	Höhere Glaubwürdigkeit; mehr parallele Opfer	HOCH
Unternehmenszugang	Imitation von Führungskräften, Dienstleistern oder Bewerbern	Zugangsdaten, Zahlungen, interne Informationen	HOCH
Öffentliche Debatte	Viele langlebige Konten mit individuellen Beziehungen	Künstlicher Konsens; schwer messbare reale Mehrheiten	MITTEL
Gefälschte Belege	Bilder, Video, Dokumente und Zeugenaussagen stützen ein Narrativ	Verzögerte Prüfung; falsche operative Entscheidungen	MITTEL
Identitätsübernahme	Stimme und Verhalten realer Personen werden nachgebildet	Vertrauenskanäle verlieren Beweiskraft	HOCH

SYSTEMISCHES RISIKO

Digitale Räume behandeln sichtbare Aktivität häufig als indirekten Hinweis auf reale Menschen: viele Stimmen, viele Konten, viele unabhängige Bestätigungen. Diese Annahme wird schwächer, sobald eine kleine Zahl von Akteuren dauerhaft eine große Zahl konsistenter Identitäten betreiben kann.

Nicht jede falsche Identität muss überzeugen. Bei sehr großer Zahl genügt eine geringe Erfolgsquote.

PRIORITÄTSSIGNALE

- Identische Gesprächslogik bei stark unterschiedlichen Rollen, Sprachen oder Zielgruppen.
- Ungewöhnlich hohe Reaktionsgeschwindigkeit über viele Konten und Zeitzonen hinweg.
- Gemeinsame technische oder verhaltensbezogene Merkmale trotz verschiedener Medien.
- Kontaktaufbau kurz nach öffentlich sichtbaren Lebensereignissen oder organisatorischen Änderungen.

04 / REAKTION

Operative Maßnahmen

Einzelne Erkennungsmodelle dürfen nicht als Wahrheitsinstanz behandelt werden. Maßnahmen müssen auch dann funktionieren, wenn synthetisches Material visuell und akustisch nicht sicher von echtem Material getrennt werden kann.

SOFORT UMSETZBAR

MASSNAHME	UMSETZUNG	ZIEL
Kanalwechsel	Hochrisikoaktionen über einen unabhängig bekannten Kanal bestätigen.	Angreifer vom kontrollierten Gespräch trennen.
Herkunft sichern	Originaldateien, Übertragungsweg, Zeitstempel und Gerätebezug früh bewahren.	Provenienz statt sichtbarer Plausibilität prüfen.
Vier-Augen-Freigabe	Neue Empfänger, hohe Zahlungen und Rechteänderungen nie nur per Stimme oder Nachricht freigeben.	Wirkung einzelner Imitationen begrenzen.
Fallübergreifend suchen	Stimmen, Formulierungen, Rhythmen, Konten, Geräte und Reaktionsmuster vergleichen.	Wiederverwendung und Infrastruktur erkennen.
Unsicherheit benennen	Ergebnisse als Wahrscheinlichkeiten dokumentieren; Rohdaten erhalten.	Scheinsicherheit durch Detektoren vermeiden.

ERSTREAKTION BEI VERDACHT

- Kommunikation sichern, ohne den mutmaßlichen Betreiber sofort auf die Untersuchung hinzuweisen.
- Keine weitere Zahlung, Rechtevergabe oder Datenweitergabe über denselben Vertrauenskanal.
- Betroffene Person oder Organisation über zuvor verifizierte Kontaktdaten erreichen.
- Material nicht nur einzeln, sondern als zusammenhängende Identität auswerten.
- Ähnliche Fälle zentral verknüpfen und Sprach-/Bildmaterial für spätere Verfahren erhalten.

NICHT AUSREICHEND

Ein Videocall, ein Ausweisdokument oder eine glaubwürdig klingende Rückfrage ist allein kein belastbarer Identitätsnachweis. Je stärker ein Verfahren auf genau einem Medium beruht, desto leichter kann dieses Medium zum Angriffspunkt werden.

05 / FORTSCHREIBUNG

Ausblick und Prüfmarker

Die technische Entwicklung ist für dieses Lagebild weniger entscheidend als ihre Zusammenführung. Das Risiko steigt sprunghaft, wenn Bild, Stimme, Text, Erinnerung und Verhalten nicht mehr getrennt erzeugt, sondern als eine dauerhaft gepflegte Identität koordiniert werden.

BEOBACHTUNGSAUFRÄGE

PRÜFMARKER	FRAGE FÜR DIE NÄCHSTE FASSUNG
Echtzeitstimme	Sind freie Gespräche ohne hörbare Verzögerung und mit stabiler Emotionalität möglich?
Multimodale Konsistenz	Bleiben Biografie, Gesicht, Stimme und Schreibstil über Monate widerspruchsfrei?
Autonome Auswahl	Können Systeme Ziele selbst priorisieren und den Kontaktzeitpunkt optimieren?
Parallele Beziehungen	Wie viele individuelle Dialoge kann ein Betreiber glaubwürdig überwachen?
Herkunftsnachweise	Welche technischen Verfahren bleiben belastbar, wenn Metadaten manipulierbar sind?

Das Problem ist nicht, dass Maschinen plötzlich böse werden.
Das Problem ist, dass menschliche Absichten billiger werden.

ARBEITSVERMERK

Fassung 1 gespeichert. Fortschreibung nach Abschluss der laufenden Laborprüfung, Abgleich weiterer Verfahren und Bewertung neuer Echtzeit-Sprachsysteme. Schwerpunkt der nächsten Fassung: Skalierung.

FREIGABE

VERFASSERIN	PRÜFUNG	VERTEILER
M. Voss	ausstehend	begrenzt / intern

ANLAGE / FALLDATEN

[REDACTED]

[REDACTED]

[REDACTED]

ENDE DER ARBEITSFASSUNG // SI-OR-2022-001