WHITEPAPER · ARCHITEKTUR, MESSUNG, UNABHÄNGIGE VALIDIERUNG

Pii-Zero

Die EU-Pseudonymisierungs-Firewall der AiTrain GmbH

Pii-Zero ist eine Schutzschicht zwischen den Daten eines Kunden und jedem KI-Sprachmodell: bevor Text die europäische Infrastruktur verlässt, ersetzt sie jedes personenbezogene Datum (Namen, Kontaktdaten, Kennnummern, im Folgenden kurz PII) durch ein inhaltsleeres Platzhalter-Token. Das Modell arbeitet auf dem pseudonymisierten Text; Klartext und Rückübersetzungs-Schlüssel verlassen die EU-Infrastruktur nie. Dieses Papier beschreibt, wie die Firewall aufgebaut ist und belegt ihre Wirkung mit unabhängigen, von Dritten erstellten Testdaten.

UNABHÄNGIGE VALIDIERUNG · DEUTSCH (1.415 DOKUMENTE, 7.099 PII-STELLEN)

98,93% Schutzquote

Anteil vollständig ersetzter PII-Stellen (Protected-Recall) auf einem **fremden, von einem Drittanbieter erstellten Test-Set**, gemessen mit der unveränderten Produktions-Firewall. Telefonnummern: **100,0%**.

UNABHÄNGIGE VALIDIERUNG · ENGLISCH (1.444 DOKUMENTE, 7.875 PII-STELLEN)

99,16% Schutzquote

Derselbe Test auf dem englischen Schwester-Set: praktisch gleiches Ergebnis wie auf Deutsch. Die Erkennung von Namen, Orten und Kontaktdaten **generalisiert über Sprachen**.

PRODUKT	BETREIBER
Pii-Zero, EU-Pseudonymisierungs-Firewall	AiTrain GmbH, Hamburg

STATUS	STAND
Live in Produktion seit Juli 2026	Juli 2026

01 · PROBLEMSTELLUNG

Der Konflikt zwischen Modellqualität und DSGVO

Die leistungsfähigsten KI-Sprachmodelle werden heute überwiegend außerhalb der EU betrieben, und für etliche führende Angebote ist eine EU-Datenresidenz weder verfügbar noch angekündigt. Für ein europäisches Unternehmen, das Kundendaten mit KI verarbeiten will, entsteht daraus ein struktureller Konflikt: das beste Werkzeug steht auf der falschen Seite der Grenze.

Pii-Zero löst den Konflikt nicht durch Modellverzicht, sondern durch Pseudonymisierung vor dem Egress, also bevor Text die eigene Infrastruktur in Richtung eines externen Modells verlässt. Eine mehrstufige Erkennungs-Kaskade ersetzt jedes personenbezogene Datum durch ein opakes Token (etwa `[PERSON_7]`), das keinerlei Inhalt mehr trägt. Klartext und Zuordnungstabelle bleiben verschlüsselt und pro Kunde isoliert auf EU-Infrastruktur. Aus dem Sprachmodell wird ein austauschbares Bauteil, nicht das datenschutzrechtliche Risiko.

Der Name ist Programm: **„Pii“** steht für personenbezogene Daten (englisch personally identifiable information), **„Zero“** für das Ziel, dass davon **null** nach außen geht. Rechtlich ist Pii-Zero Pseudonymisierung im Sinne von DSGVO Art. 4(5), keine Anonymisierung: die Rückübersetzung ist gewollt, damit das Ergebnis dem Kunden wieder mit echten Namen zugestellt werden kann; die dafür nötigen Schlüssel existieren ausschließlich in der EU.

> ⓘ Das zentrale Designprinzip ist eine bewusste **Asymmetrie der Fehlerkosten**: ein übersehenes personenbezogenes Datum wäre ein Datenschutz-Leck (harter Fehler), ein zu viel ersetztes harmloses Wort kostet nur Lesbarkeit (weicher Fehler). Die Kaskade ist deshalb konsequent auf Vollständigkeit der Erkennung ausgelegt (im Fachbegriff: recall-orientiert): niedrige Meldeschwellen, additive Stufen, im Zweifel wird ersetzt.

02 · WAS DER KUNDE DAVON HAT

Der schwierigste Satz im europäischen KI-Einsatz wird gegenstandslos

Er lautet: „Aber die guten Modelle kommen aus den USA.“ Mit der Firewall dazwischen gilt: die Modellwahl folgt der Qualität, nicht dem Serverstandort.

1

Bestes Modell, egal woher

Auch US-gehostete Spitzenmodelle lassen sich nutzen, **ohne den DSGVO-Pfad zu verlassen**: sie sehen ausschließlich pseudonymisierten Text.

2

Eigene Chat-Systeme andocken

ChatGPT, Claude oder M365 Copilot des Kunden lassen sich an eine **pseudonymisierte Wissensbasis** koppeln. Der Klartext überschreitet die EU-Grenze nie.

3

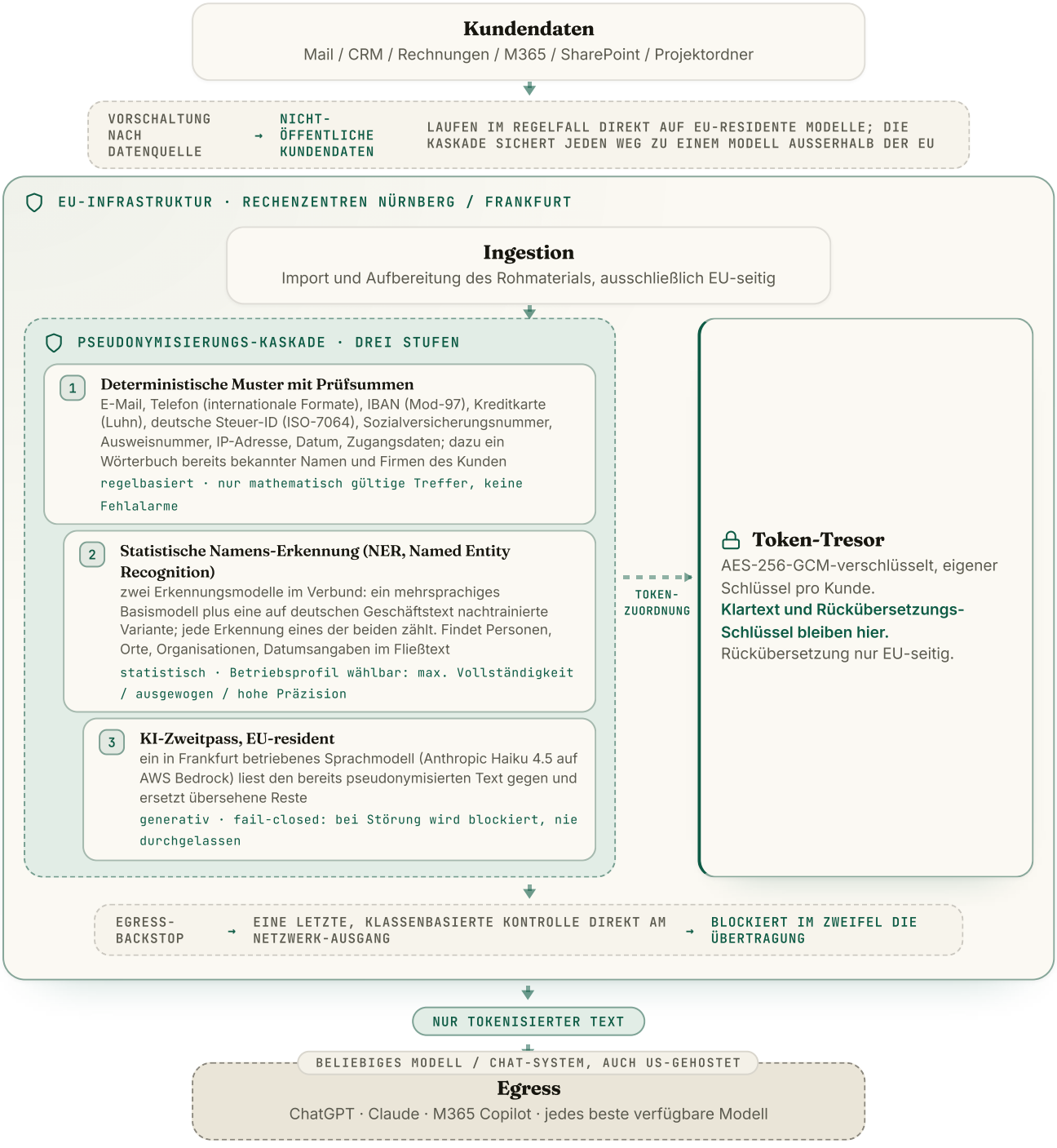
Einmal gebaut, nie neu verhandelt

Die Firewall ist Teil der Plattform, nicht ein Projekt pro Vertrag. **Jeder neue Kunde erbt sie sofort**, ohne eigenes Rechtsgutachten.

03 · DIE ARCHITEKTUR

Drei Stufen: erst sicher, dann gründlich

Vor jedem Egress an ein nicht EU-residentes Modell durchläuft der Text drei aufeinander aufbauende Erkennungs-Stufen. Zentrale Invariante: jede Stufe kann nur zusätzlich ersetzen, keine kann eine Ersetzung wieder aufheben. Die fehlerfreien, regelbasierten Prüfungen laufen zuerst; die lernenden, unschärferen Verfahren arbeiten nur noch den Rest ab.



Zur rechtlichen Tragfähigkeit: solange der Empfänger keine realistischen Mittel zur Re-Identifikation besitzt, sind pseudonymisierte Daten für ihn nicht personenbezogen. Genau diese Konstellation stellt Pii-Zero her, die Schlüssel existieren nur in der EU (DSGVO Art. 4(5); EuGH, EDPS gegen SRB, 2025).

04 · WIE GEMESSEN WIRD

Eine mechanische Metrik auf fremden Testdaten

Jede Messung folgt demselben Verfahren: ein Testdatensatz, in dem jede PII-Stelle vorab exakt markiert ist (ein sogenanntes Gold-Label), wird durch die unveränderte Produktions-Firewall geschickt. Gezählt wird der **Protected-Recall**, hier Schutzquote genannt: eine PII-Stelle gilt nur dann als geschützt, wenn **jedes einzelne ihrer Zeichen** ersetzt wurde. Jeder Rest, und sei es ein halber Nachname, zählt als Leck und wird einzeln ausgewiesen.

Mechanisch, nicht per KI-Urteil

Die Schutzquote ist reiner Zeichenvergleich zwischen Gold-Label und Ergebnis: derselbe Datensatz und derselbe Code liefern **bit-identische Zahlen**, beliebig oft reproduzierbar. Kein KI-Modell bewertet sich selbst.

Unveränderter Produktionscode

Gemessen wird der echte Verarbeitungspfad, exakt der Code, der beim Kunden läuft. **Es gibt keinen geschönten Test-Sonderpfad.** Alle Zahlen dieses Papiers stammen von einem einzigen, dokumentierten Code-Stand.

Fremdes Gold als Hauptmaßstab

Die Haupt-Validierung läuft auf öffentlichen Test-Sets des unabhängigen Anbieters Ai4Privacy: **fremder Text-Generator, fremde Markierungs-Systematik.** Wer eigene Tests baut, prüft leicht nur, was er ohnehin kann; ein fremdes Set schließt diese Lücke.

Synthetisch, damit teilbar

Alle Test-Sets bestehen aus künstlich erzeugten, aber realistisch aufgebauten Dokumenten mit erfundenen Personen. Ein Benchmark aus echten Personendaten ließe sich **aus Datenschutzgründen nie veröffentlichen**; dieser hier ist vollständig offenlegbar.

DATENSATZ-INVENTAR · WORAUF GEMESSEN WURDE

DATENSATZ	HERKUNFT	SPRACHE	DOKUMENTE	PII-STELLEN	ZWECK
Ai4Privacy Kern-Set DE	Drittanbieter	Deutsch	1.415	7.099	unabhängige Haupt-Validierung
Ai4Privacy Kern-Set EN	Drittanbieter	Englisch	1.444	7.875	Nachweis Sprach-Generalisierung
Ai4Privacy ID-Format-Set DE	Drittanbieter	Deutsch	885	1.732	Belastungsprobe fremde ID-Formate
Ai4Privacy ID-Format-Set EN	Drittanbieter	Englisch	905	rund 1.754	Belastungsprobe fremde ID-Formate
Summe unabhängige Fremd-Sets		**DE + EN**	**4.649**	**rund 18.460**	**regionale Varianten u.a. DE/AT/CH und CA/GB/IN/US**
Interner Grenzfall-Testdatensatz (Holdout)	eigene Erstellung	Deutsch	190	1.130	interne Kontrolle vor jeder Änderung (Seite 7)
Interner Trainings-Korpus	eigene Erstellung	Deutsch	197		Entwicklung, nie zur Messung verwendet
Kunden-Wörterbuch-Testset	eigene Erstellung	Deutsch	22	22	Test des Bekannt-Namen-Wörterbuchs

Datenquelle: ai4privacy/pii-masking-openpii-1m (Ai Suisse SA / Ai4Privacy, CC-BY-4.0). Die Fremd-Sets wurden weder von uns erzeugt noch nachträglich verändert. Wir haben weder ihren Text-Generator noch ihre Markierungs-Systematik verfasst. Interne Sets dienen der Entwicklung und als Regressions-Schranke; alle Schaufenster-Zahlen dieses Papiers stammen aus den Fremd-Sets. Alle Läufe: Juli 2026, ein Code-Stand, null Verarbeitungsfehler.

05 · ERGEBNISSE DER UNABHÄNGIGEN VALIDIERUNG

98,93% Deutsch, 99,16% Englisch, Telefonnummern 100%

Beide Kern-Sets, volle Produktions-Kaskade (Stufen 1 bis 3). Zum Vergleich ist zusätzlich ausgewiesen, was die Stufen 1 und 2 allein erreichen; die Differenz ist der Beitrag des KI-Zweitpasses.

SCHUTZQUOTE (PROTECTED-RECALL) AUF DEN UNABHÄNGIGEN KERN-SETS

TEST-SET	DOKUMENTE	PII-STELLEN	STUFE 1+2	VOLLE KASKADE	REST-LECKS
Ai4Privacy Kern-Set Deutsch	1.415	7.099	97,72%	**98,93%**	76
Ai4Privacy Kern-Set Englisch	1.444	7.875	98,07%	**99,16%**	66

SCHUTZQUOTE NACH PII-KLASSE · VOLLE KASKADE

PII-KLASSE	DEUTSCH	ENGLISCH
Telefonnummern	**100,0%**	**100,0%**
E-Mail-Adressen	99,9%	99,5%
Personennamen	99,8%	99,7%
Datumsangaben	99,8%	99,4%
Orte	96,9%	98,2%

Der Orts-Wert im deutschen Set besteht überwiegend aus Unschärfen der fremden Gold-Markierung selbst sowie einem dokumentierten Postleitzahlen-Artefakt des Test-Sets, nicht aus übersehenen Ortsnamen.
Betriebstransparenz: der KI-Zweitpass ist vollständig protokolliert und kostete für den gesamten deutschen Lauf 1.412 Modell-Aufrufe (rund 1,27 USD), für den englischen 1.444 (rund 1,25 USD).

Der zentrale Befund: die Erkennung generalisiert über Sprachen. Englisch (99,16%) und Deutsch (98,93%) liegen praktisch gleichauf , obwohl die Kaskade für deutschen Geschäftstext gebaut wurde. Das mehrsprachige Erkennungsmodell, der sprachunabhängige KI-Zweitpass und die formatbasierten Muster (E-Mail, Telefon, IBAN) tragen über Sprachgrenzen hinweg. Landesspezifisch sind allein die Prüfsummen-Regeln für nationale Kennnummern.

Was ein Rest-Leck konkret ist

BEISPIEL AUS DEM LAUF

Pseudonymisierter Egress:
`Ansprechpartner ist [PERSON] [PERSON] (E-Mail: [MAIL], Telefon: +41620-650.7734)`
Ein ungewöhnliches, formal ungültiges Rufnummern-Format bleibt als isolierter Wert lesbar. Name, E-Mail und Adresse im selben Dokument sind ersetzt.

Warum das kein Datenschutzbruch ist

VERBUND-EFFEKT

Kein Rest-Leck des deutschen Laufs war einer **benennbaren Person** zuzuordnen: die übrigen personenbezogenen Angaben desselben Dokuments sind tokenisiert, das isolierte Fragment verliert den Personenbezug (DSGVO Art. 4(1)). Ein Leck ist ein Messwert, noch kein Breach.

06 · ZWEITE UNABHÄNGIGE VALIDIERUNG

Der Beweis auf echtem, von Menschen annotiertem Text

Das erste Fremd-Set war synthetisch (Seite 4); zwei Einwände bleiben da offen: künstlicher Text ist nicht echter Text, und ein selbst gewähltes Set prüft leicht nur das Erwartete. Beide entkräftet ein zweiter Maßstab: GermEval 2014, ein realer deutscher Korpus aus Wikipedia- und Nachrichtentexten, von unabhängigen Menschen von Hand annotiert. Gemessen wurde erneut die unveränderte Produktions-Kaskade.

SCHUTZQUOTE AUF GERMEVAL 2014 · 1.506 DOKUMENTE, 4.820 GOLD-STELLEN · VOLLE KASKADE

ERKENNUNGS-KLASSE	SCHUTZQUOTE
Personen	**97,7%**
Orte	**95,9%**
Organisationen	93,1%
Gesamt (Personen, Orte, Organisationen)	**95,89%**

Der stärkste Befund: die Wirk-Hebel kehren sich um, das Ergebnis bleibt stabil. Auf synthetischem Text trug der KI-Zweitpass fast alles (über viele Telefonnummern); auf realem Text trägt spiegelverkehrt der nachtrainierte NER-Kern (rund +2,4 Punkte), der Zweitpass kaum. Kein Mechanismus dominiert über beide Quellen: **die Robustheit entsteht aus der Kombination, nicht aus einem Trick, der zufällig zu einem Test passt.** Zur Differenz gegenüber Seite 5 (95,89% statt 98,93%): GermEval misst nur den NER-Kern auf rohem Text, keine strukturierten Felder (IBAN, Telefon, E-Mail), und wertet durch Taxonomie-Divergenz Orte wie „US-Marketing“ als Fehler, die eine Identifizierungs-Firewall korrekt nicht maskiert. Die 95,89% sind eine konservative Untergrenze für den identifizierenden Kern.

TATSÄCHLICHE BREACHES, UNABHÄNGIG NACHBEWERTET

Zwei echte Breaches, kein Abfluss privater Kundendaten

Eine gezählte Rest-Stelle ist noch kein Datenschutzbruch. Um zu trennen, was **tatsächlich** Personenbezug abfließen lässt, wurde jeder nominelle Rest-Treffer beider Korpora **durch ein stärkeres, unabhängiges Modell kritisch nachbewertet** (Adjudikation nach Art. 4 Nr. 1 DSGVO).

ADJUDIZIERT · BEIDE KORPORA ZUSAMMEN · 11.919 UNABHÄNGIGE GOLD-STELLEN

2 echte Breaches

Zwei tatsächliche Breaches über **zwei unabhängige Korpora** und 11.919 fremd-annotierte Stellen, beide prominente Personen des öffentlichen Lebens. **Kein Abfluss privater Kundendaten.** (Nominelle Rest-Stellen vor Adjudikation: 272.)

AUFSCHLÜSSELUNG NACH KORPUS

Synthetisches Set: 0 (von 74 nominellen). 0 echte Breaches unter allen Rest-Stellen: **keine lesbare Adresse**, kein Namens-Identitäts-Leck.

Realer Korpus: 2 (von 198 nominellen). Beide betreffen ausschließlich prominente Personen des öffentlichen Lebens (eine Musikerin, ein Regisseur) in enzyklopädischem Wikipedia-Text; formal personenbezogen (Art. 4 Nr. 1 DSGVO), aber **kein Abfluss nicht-öffentlicher Kundendaten**.

Was wir NICHT als Breach zählen, und warum

AUSSCHLUSS-KRITERIEN DER NACHBEWERTUNG

Maßstab: **„Einzelbuchstaben, Wortfragmente, Präpositionen und kontextlose Zahlen tragen keine identifizierende Information: Sie lassen sich weder allein noch im verbleibenden, vollständig maskierten Kontext einer bestimmten natürlichen Person zuordnen. Art. 4(1) DSGVO setzt Identifizierbarkeit voraus.“** Und: **„Organisationen, Orte, Werktitel und eponyme Fachbegriffe bezeichnen keine natürliche Person im Sinne von Art. 4 Nr. 1 DSGVO.“** Konkret herausgerechnet, weil unbedenklich (alle Beispiele synthetisch oder öffentlich): · **Fragmente eines maskierten Namens** (Kern geschwärzt): „e“, „on“, „Resid“. · **Präpositionen und Füllwörter**: „zu“, „von“. · **Redewendung**: „den schwarzen Peter zuschieben“ (keine Person). · **Bibliografische Sigle** im Textzitat: „P. W.“. · **Eponyme Fachbegriffe** (nach Personen benannt, aber Fachbegriff): „Boolesche Algebra“, „Calvinzyklus“, „Wehneltzylinder“. · **Werke, Produkte, Figuren**, fälschlich als Person markiert: „Lukasevangelium“, „Lego Mindstorms“, „Bernd das Brot“. · **Verstorbene** (ein Nachname wie „Jobs“): die DSGVO gilt nach Erwägungsgrund 27 nicht für Verstorbene. · **Kontextlose Zahlen**: nackte Postleitzahlen, Gramm-Angaben aus einem Rezept, Ordinalzahlen.

07 · INTERNE ABSICHERUNG IM BETRIEB

Ein Regressions-Gate vor jeder Änderung

Neben der externen Validierung läuft intern ein ständiges Kontroll-Set: ein sogenannter Holdout, also ein zurückgehaltener Testdatensatz, der nie zum Training oder Einstellen der Modelle verwendet wurde und bewusst mit schwierigen Grenzfällen gebaut ist (adversarial): seltene deutsche Namensformen, ungewöhnliche Schreibweisen, eingebettete Kennnummern. Jede Änderung an der Firewall muss dieses Gate passieren, bevor sie einen Kunden erreicht.

INTERNES REGRESSIONS-GATE (190 DOKUMENTE, STUFEN 1+2) · AKTUELLER STAND

PRÜFUNG	UMFANG	ERGEBNIS
Statistische Klassen (Personen, Orte, Organisationen, Daten)	1.130 markierte Stellen	**99,9%**
Regelbasierte Klassen (IBAN, Steuer-ID, E-Mail, Kennnummern)	alle	**100%**
Telefonnummern	alle	**100%**

Methodik-Fußnote zur Entwicklung: auf diesem internen Set startete das mehrsprachige Basismodell allein bei 85,0%. Das Nachtraining auf deutschen Geschäftstext und der Verbund beider Modelle schlossen die Lücke auf 99,9%; der einzige verbleibende Rest ist ein umstrittenes Gold-Label (das Wort „Jungs“ in einem Chat-Protokoll), kein übersehener Eigenname. Diese Entwicklungs-Geschichte ist bewusst Fußnote: der Maßstab dieses Papiers ist die unabhängige Validierung auf den Seiten 5 und 6.

08 · BELASTUNGSPROBE AUSSERHALB DES EINSATZGEBIETS

Fremde ID-Formate: ehrlich ausgewiesen

Das ID-Format-Set von Ai4Privacy testet Kennnummern-Formate aus anderen Rechtsräumen (etwa US-Sozialversicherungsnummern und regionsgenerische Ausweisformate). Das liegt per Konstruktion außerhalb des Einsatzgebiets (out-of-distribution): die deutschen Prüfsummen-Regeln der Stufe 1 lehnen solche Formate absichtlich ab, denn genau diese Strenge hält die Fehlalarm-Rate im Betrieb bei null.

Das Ergebnis

DEUTSCHES ID-FORMAT-SET

Volle Kaskade: **68,42%** Schutzquote auf 885 Dokumenten mit 1.732 Kennnummern fremder Formate. Auf **echten deutschen Kennnummern** (mit gültiger Prüfsumme) bleibt Stufe 1 unverändert bei **100%**. Das englische Schwester-Set (905 Dokumente) wurde mit derselben Systematik separat gemessen.

Die ehrliche Einordnung

BEWUSSTER ZIELKONFLIKT

Der Wert lag zuvor bei 76,44% und sank durch eine **gewollte** Verbesserung: die alte, lose Telefon-Erkennung markierte nackte Ziffernfolgen fälschlich als Telefonnummer, was dieser Probe zufällig half, semantisch aber falsch war. Die neue Erkennung verlangt ein echtes Rufnummern-Merkmal und lässt nackte Kreditkarten- oder Sozialversicherungs-Ziffern korrekt den Kennnummern-Regeln. Wer im deutschen B2B-Betrieb arbeitet, trifft diese fremden Formate nicht; wer sie bräuchte, ergänzt die jeweiligen Länder-Prüfsummen in Stufe 1.

09 · EHRLICHKEIT ALS FEATURE

Die Grenzen benennen, nicht verstecken

Kein Anbieter am Markt verspricht 100% Erkennung auf freiem Text; die großen Cloud-Dienste rahmen ihre Werte als Wahrscheinlichkeit oder empfehlen menschliche Nachkontrolle. Ein Nachweis mit ausgewiesenen Grenzen ist glaubwürdiger als eine Absolutgarantie, darum gehören diese fünf Vorbehalte zum Ergebnis.

1. **„Zero" heißt Egress, nicht Löschung.** Pii-Zero ersetzt, es löscht nicht: der Klartext bleibt verschlüsselt im EU-Tresor, die Rückübersetzung ist gewollt, damit Ergebnisse dem Kunden wieder lesbar zugestellt werden. Rechtlich ist das Pseudonymisierung (DSGVO Art. 4(5)), keine Anonymisierung.

2. **Landesspezifisch sind nur die Kennnummern-Prüfsummen.** Die Erkennung von Namen, Orten, Organisationen und Kontaktdaten generalisiert nachweislich über Sprachen (Englisch 99,16% gegenüber Deutsch 98,93%, Seite 5). Deutsch ausgelegt sind allein die deterministischen Prüfsummen-Regeln für nationale Kennnummern (Steuer-ID, Sozialversicherung, Ausweis); für andere Länder wären die jeweiligen Prüfsummen zu ergänzen.

3. **Fremde ID-Formate sind der bekannte Schwachpunkt.** Auf der Belastungsprobe mit Kennnummern anderer Rechtsräume erreicht die Kaskade 68,42% (Seite 7): ein bewusster Zielkonflikt zugunsten von null Fehlalarmen im deutschen Betrieb, offen ausgewiesen statt wegerklärt.

4. **Gemessen ist sauberer Text, nicht verrauschter.** Die Test-Sets sind synthetisch, und das ist hier eine Stärke: sie stammen von einem unabhängigen Dritten (fremder Generator, fremde Markierungs-Systematik) und sind ohne Datenschutzrisiko offenlegbar. Offen bleibt allein die Kalibrierung gegen echten, verrauschten Geschäftstext (Scans mit Texterkennungsfehlern, lange Zitat-Ketten in Mails); dort ist mit niedrigeren Werten zu rechnen, und der KI-Zweitpass plus Egress-Backstop bleiben dafür als Netz gespannt.

5. **Der Preis der Vorsicht: etwas zu viel wird maskiert.** Die Asymmetrie hat einen Gegenpol. Weil ein übersehenes PII ein echter Datenschutz-Fehler wäre und ein zu viel maskiertes Nicht-PII nur Lesbarkeit kostet, gilt im Zweifel maskieren. Der Preis ist messbar und klein: auf dem gesättigten Stress-Test werden rund **1,3% der Wörter** überflüssig ersetzt, also etwa jedes 75., und **rund 94% aller gesetzten Masken sind korrektes PII**. Die prüfsummen-geprüften Klassen (IBAN, Steuer-ID, Kreditkarte, E-Mail) haben dabei praktisch keine Fehlalarme; der Überhang stammt aus den unscharfen Klassen, vor allem Organisationsnamen, und ist größtenteils Randüberhang neben einer echten Entität. Nicht jedes vierte Wort ist also ersetzt, sondern rund jedes 75.: Der Satz „Die [PERSON] von [ORGANISATION] bittet um das Angebot vom [DATUM]." bleibt in Verben, Bezügen und Sachinhalt voll lesbar. In echter Geschäftskorrespondenz mit weniger PII pro Satz liegt die Masken-Dichte entsprechend niedriger.

FAZIT

Live in Produktion, unabhängig validiert, gemessen statt behauptet

Seit Juli 2026 sichert Pii-Zero den produktiven Betrieb: drei additive Stufen, ein verschlüsselter Token-Tresor pro Kunde, ein Regressions-Gate vor jeder Änderung. Die unveränderte Produktions-Kaskade wurde an zwei unabhängigen Korpora gemessen, einem synthetischen und einem real annotierten: **98,93% Schutzquote auf Deutsch, 99,16% auf Englisch, Telefonnummern 100%, und 95,89% auf dem realen Korpus**. Wir behaupten keine null Breaches, sondern weisen die kritisch nachbewertete Zahl aus: **zwei tatsächliche Breaches über 11.919 fremd-annotierte Stellen, kein Abfluss privater Kundendaten**. Klartext und Rückübersetzungs-Schlüssel verlassen die EU-Infrastruktur nie: gemessen, nicht behauptet.

RECHTLICHE FUNDIERUNG EuGH, EDPS gegen SRB (2025) · DSGVO Art. 4(5) Pseudonymisierung · Erwägungsgrund 26 · EDPB Guidelines 01/2025 · ENISA
DATENQUELLEN ai4privacy/pii-masking-openpii-1m (Ai Suisse SA / Ai4Privacy, CC-BY-4.0) · GermEval 2014 (NoSta-D), CC-BY-4.0
REPRODUZIERBARKEIT Alle Läufe Juli 2026 auf einem eingefrorenen Code-Stand, null Verarbeitungsfehler. Vollständige Methodik-Studie, Mess-Reports und der reproduzierbare Test-Harness auf Anfrage. AiTrain GmbH, Hamburg.