

BENUTZERDOKUMENTATION

Sovereign Cockpit

Das Handbuch für den täglichen Gebrauch: anonymisieren, prüfen, freigeben, re-identifizieren, und die Verwaltung von Standardbegriffen, Vorgängen und Benutzerkonten.

Inhalt

- | | |
|--------------------------------|--|
| 1. Das Grundprinzip | 8. Freigeben und herunterladen |
| 2. Anmelden, Sprache, Abmelden | 9. Antwort re-identifizieren |
| 3. Dokument anonymisieren | 10. Vorgänge verwalten |
| 4. Erkannte Begriffe prüfen | 11. Benutzer und Rollen (Administration) |
| 5. Die Prüfseite | 12. Sicherheit und Datenhaltung |
| 6. Manuell nachmaskieren | 13. Grenzen und Tipps |
| 7. Standardbegriffe | 14. Meldungen und Hilfe |

1 Das Grundprinzip

Sovereign Cockpit sitzt **vor** jedem Versand nach draussen: an ein Cloud-KI-Modell, per E-Mail oder als Weitergabe. Vertrauliche Daten werden erkannt und durch Platzhalter wie `<PERSON_1>` oder `<AHV_1>` ersetzt. Nur die entschärfte Fassung verlässt das System, und nur durch Ihren ausdrücklichen Klick im **Human-Gate**. Die Zuordnung Platzhalter ↔ Original (das **Mapping**) bleibt verschlüsselt auf dem Server; damit setzt das Cockpit die Originalwerte in der Antwort des KI-Modells wieder ein.

Drei Grundsätze prägen die Bedienung:

- **Sie entscheiden.** Jede Ersetzung ist sichtbar und abwählbar; jede Auslieferung braucht Ihren Klick.
- **Fail-closed.** Scheitert ein Verarbeitungsschritt, bricht der Vorgang mit einer klaren Meldung ab: es werden nie stillschweigend weniger Daten geschützt.
- **Nachvollziehbar.** Jeder Schritt (Anonymisierung, Freigabe, Löschung) wird in einem manipulationssicheren Audit-Protokoll festgehalten.

2 Anmelden, Sprache, Abmelden

Sie erreichen Ihr Cockpit über die Adresse Ihrer Instanz (z.B. `https://<ihre-instanz>.iconference.ch`). Melden Sie sich mit Benutzername und Passwort an; die Konten werden von Ihrer Administratorin/Ihrem Administrator angelegt (Kapitel 11).

- **Sprache:** oben rechts zwischen DE und EN umschalten, jederzeit, ohne Datenverlust.
- **Passwort ändern:** unter **Profil** (mindestens 10 Zeichen).
- **Abmelden:** oben rechts. Nach längerer Inaktivität endet die Sitzung automatisch.

3 Dokument anonymisieren

Auf der Startseite **Anonymisieren** haben Sie zwei Wege:

- **Text einfügen:** direkt ins Textfeld (bis 150'000 Zeichen).
- **Datei hochladen:** PDF (mit Textebene), Word (.docx), Markdown (.md), Text (.txt) oder Excel (.xlsx), bis 25 MB.

Die **Bezeichnung** ist optional und dient der Wiedererkennung in der Vorgangsliste; leer gelassen wird der Dateiname (oder «Direkteingabe») plus Datum verwendet.

Gescannte PDFs werden abgelehnt. Ein PDF ohne Textebene (Foto/Scan) kann nicht zuverlässig geschützt werden: das Cockpit bricht bewusst ab, statt unsicher zu raten. Wandeln Sie das Dokument zuerst mit einer OCR-Lösung in durchsuchbaren Text um oder nutzen Sie die Textquelle.

- **Excel:** Zellen, Blattnamen, Kommentare, Kopf-/Fusszeilen und Text-Literale in Formeln werden geprüft; Zahlen und Formelstruktur bleiben unangetastet. Postleitzahl- und Geburtsdatum-Spalten werden am Spaltentitel erkannt (z.B. «PLZ», «Geburtsdatum») und als maskierbare Werte angeboten, obwohl Excel sie als Zahl bzw. Datum speichert.
- **Markdown:** läuft formatiert durch: Titel, Listen und Auszeichnungen bleiben erhalten, und die Downloads tragen wieder die Endung .md.
- **Word (.docx):** geprüft werden Absätze, Tabellen, Kopf- und Fusszeilen, Fussnoten, Endnoten, Kommentare und Textboxen; Layout und Formatierung bleiben erhalten, und die Downloads tragen wieder die Endung .docx. Abgelehnt werden Dokumente mit aktiven Änderungsnachverfolgungen (Tracked Changes), eingebetteten Objekten sowie passwortgeschützte oder alte .doc-Dateien. Heben Sie die Nachverfolgung auf bzw. speichern Sie als .docx.

Während der Verarbeitung zeigt das Cockpit eine laufende Fortschrittsanzeige mit Sekundenzähler. Braucht der erste Start der Instanz länger (Kaltstart), weist ein eigener Hinweistext darauf hin.

4 Erkannte Begriffe prüfen

Nach dem Start zeigt das Cockpit den Originaltext mit allen erkannten Stellen. Eine Checkliste oben fasst zusammen, wie viele Stellen automatisch maskiert sind und wie viele Firmen-, Personen-, Geburtsdaten- und Rollen-Treffer zur Prüfung anstehen. Die Vorschläge sind in vier Bereiche gegliedert:

- **Automatisch maskiert.** Eine zugeklappte Sammel-Sektion für eindeutige Identifikatoren (AHV, IBAN, Firmen-UID, E-Mail/Telefon, Aktenzeichen, Grundstück-Kennungen, Adressen ...), alle vorgehakt; die Kopfzeile zeigt die Anzahl je Typ.
- **Firmen- und Personen-Karten.** Eine Karte pro Firma oder Person: alle im Dokument gefundenen Schreibweisen (Dr., Herr, Initialen, Kurzformen ...) erscheinen darauf als einzeln abwählbare, vorgehakte Chips. Ein erkanntes Geburtsdatum erscheint ebenfalls als vorgehakter Chip direkt auf der Personen-Karte. Alle Schreibweisen derselben Person oder Firma laufen auf denselben Platzhalter. Die Re-Identifikation setzt danach durchgehend die vollständige Form ein, der anonymisierte Text bleibt lesbar und konsistent.
- **Rollen.** Vertragsrollen (etwa Verkäuferin, Willensvollstrecker) erscheinen als reine Information. Sie werden nicht maskiert und sind nicht Teil der Prüfentscheidung.

- **Daten & Orte / Sonstige.** Zwei aufklappbare Rest-Gruppen: «Daten & Orte» bündelt Datumsangaben, Ortsnamen sowie PLZ/Ort-Angaben an einer Stelle; «Sonstige» fasst den Rest. Bei den Orten stehen nur echte Ortsnamen, gegen das amtliche Ortschaftsverzeichnis geprüft; generische Fach-, Rollen- und Dokumentbegriffe (etwa GmbH, Jahresgewinn, Rechtsanwalt, Mietvertrag), welche die Erkennung fälschlich für einen Ort, eine Person oder eine Firma hielt, landen in «Sonstige». Beide Gruppen sind nie vorausgewählt.

Fachvokabular aus Rechts- und Fachgebieten (z.B. M&A, Erbrecht, Grundbuch, Bau- und Finanzwesen) erscheint dank dieser Gliederung deutlich seltener als Prüf-Kandidat. Die fachliche Prüfung bleibt bei Ihnen, der Aufwand dafür sinkt spürbar.

Fehlt ein Begriff in den Vorschlägen, tragen Sie ihn direkt hier ins Feld «Fehlt ein Begriff?» ein: Das Cockpit maskiert ihn inklusive Beugungen und Wortzusammensetzungen (Beispiel: «WEKO» deckt auch «WEKO-Vorabklärung» ab).

- **Sehr kurze Treffer** (Namen, Firmen und Orte mit zwei oder weniger Zeichen, etwa Initialen oder Kürzel) werden nicht als Vorschlag angeboten. Solche Kürzel identifizieren für sich genommen niemanden und würden die Liste nur füllen. Möchten Sie ein bestimmtes Kürzel dennoch maskieren, tragen Sie es als Standardbegriff ein oder maskieren Sie es manuell nach.

Was Sie nicht anhaken, bleibt im Klartext und verlässt das System unmaskiert. Im Zweifel anhaken: auf der Prüfseite sehen Sie das Ergebnis und können jede Entscheidung noch korrigieren.

Ihre Standardbegriffe werden hier automatisch gefunden und sind immer vorausgewählt.

5 Die Prüfseite

Nach der Bestätigung entsteht ein **Vorgang** mit einer eindeutigen Nummer. Die Prüfseite zeigt:

- den **anonymisierten Text** mit hervorgehobenen Platzhaltern,
- das **Mapping** (Original ↔ Platzhalter ↔ Typ): es bleibt serverseitig und ist nicht Teil der anonymisierten Fassung,
- das Werkzeug zum Nachmaskieren und die Download-Knöpfe (Human-Gate).

Lesen Sie den anonymisierten Text gegen, bevor Sie ihn herausgeben. Geprüft heisst nicht garantiert vollständig: Ersetzt wird, was erkannt und angehakt wurde.

Verlassen Sie die Seite über die Navigation, fragt das Cockpit nach. Der Vorgang geht dabei nicht verloren, er bleibt unter **Vorgänge** erreichbar.

6 Manuell nachmaskieren

Das Sicherheitsnetz für alles, was die Erkennung verpasst hat oder was Sie im Gate nicht angehakt haben: Tippen Sie den Begriff ins Feld **«Begriffe manuell maskieren»** und klicken Sie «Maskieren». Es ist dasselbe Feld wie «Fehlt ein Begriff?» im Gate (Kapitel 4), inklusive Beugungen und Wortzusammensetzungen.

- **Mehrere Begriffe** auf einmal: mit Komma oder pro Zeile trennen, jeder erhält eigene Platzhalter.
- **Alle Vorkommen** werden ersetzt, unabhängig von Gross-/Kleinschreibung; auch ss/ß-Varianten («Strasse» findet «Straße» und «STRASSE») und über Zeilenumbrüche zerrissene Mehrwort-Begriffe.

- **Schreibweisen-treu:** jede im Text gefundene Schreibweise erhält ihren eigenen Mapping-Eintrag: die Re-Identifikation stellt sie exakt wieder her.
- **Klare Rückmeldung:** Begriffe, die im Text nicht vorkommen, meldet das Cockpit deutlich («Nicht im Text gefunden: ...»), nichts verschwindet stillschweigend.
- **«als Standard merken»:** nimmt die eingegebenen Begriffe zusätzlich in die Standardbegriffe auf. Sie wirken ab der nächsten Anonymisierung, auch wenn sie im aktuellen Dokument nicht vorkamen.

7 Standardbegriffe

Standardbegriffe sind Ihre persönliche «immer anonymisieren»-Liste der Instanz: der eigene Firmenname, Kundennamen, Projekt- und Produktcodes. Sie werden in **jedem künftigen Dokument** automatisch gefunden (unabhängig von Gross-/Kleinschreibung, auch zusammengeschrieben, z.B. findet «UCP» auch «UCPlattform») und sind im Gate vorausgewählt.

- **Erfassen:** auf der Seite **Standardbegriffe**, mehrere Begriffe mit Komma oder pro Zeile. Oder direkt beim Nachmaskieren über den Haken «als Standard merken».
- **Entfernen:** in der Liste per Klick. Bereits maskierte Vorgänge bleiben unverändert.
- Die Liste ist verschlüsselt gespeichert und gilt für die ganze Instanz.

8 Freigeben und herunterladen (Human-Gate)

Erst mit dem Klick auf einen Download-Knopf verlässt die anonymisierte Fassung den Server. Jede Auslieferung wird mit Prüfsumme im Audit-Protokoll gesiegelt.

FORMAT	WOFÜR
.txt / .md	Empfohlen für den Rückweg (Re-Identifikation) bei Text-Antworten. Markdown-Quellen werden als .md ausgeliefert, sonst .txt.
PDF	Schlichtes Versanddokument mit Hinweis-Fusszeile, nicht für die Re-Identifikation gedacht.
Excel	Bei Excel-Quellen: die anonymisierte Arbeitsmappe, zellgenau, Zahlen und Formeln unangetastet.
Word	Bei Word-Quellen: das anonymisierte Dokument als .docx, mit erhaltenem Layout: Absätze, Tabellen, Kopf- und Fusszeilen.
Mapping (CSV)	Vertraulich: enthält die Re-Identifikations-Schlüssel im Klartext. Nur sicher aufbewahren, nicht weitergeben. Der Download fragt zur Sicherheit nach.

Restrisiko-Warnung: Vor der Auslieferung prüft eine unabhängige Egress-Kontrolle die fertige Fassung erneut. Findet sie noch kritische Werte im Klartext (z.B. AHV, IBAN, CHE-UID, Kontaktdaten oder einen Ihrer Standardbegriffe), hält der Download an und listet die betroffenen Werte einzeln mit Kontext. Sie können jeden Wert direkt dort nachmaskieren («maskieren» oder «Alle maskieren») oder das Dokument bewusst so ausliefern; eine bewusste Freigabe wird protokolliert.

9 Antwort re-identifizieren

Die Antwort des externen KI-Modells enthält noch die Platzhalter. Unter **Re-identifizieren** fügen Sie sie ein, als Text oder als Datei (PDF, Word, Markdown, Text, Excel), zusammen mit der Vorgang-Nummer (sie

ist vorausgefüllt, wenn Sie von der Prüfseite kommen).

- Das Cockpit ersetzt jeden Platzhalter durch den Originalwert, schreibweisen-treu.
- **Excel-Antworten** (z.B. die Auswertung des Modells als Arbeitsmappe) werden **zellgenau** re-identifiziert und kommen als Excel-Datei zurück: vom Modell ergänzte Spalten bleiben erhalten, und Postleitzahlen und Geburtsdaten kommen wieder als echte Zahlen- bzw. Datums-Zellen zurück, sofern sie an ihrer ursprünglichen Position stehen.
- **Word-Antworten** laden Sie als .docx hoch und erhalten sie **strukturerhaltend** re-identifiziert als Word-Datei zurück.
- **Unbekannte Platzhalter** (z.B. vom Modell verändert) werden rot markiert statt geraten.
- Download der re-identifizierten Fassung als .txt/.md, PDF, Excel oder Word.

Bereits re-identifizierten Vorgang öffnen: Ein Vorgang, der schon re-identifiziert wurde, lässt sich aus der Vorgangsliste direkt öffnen und zeigt sofort die re-identifizierte Fassung zum Download, ohne erneute Eingabe.

Danach ist der Vorgang abgeschlossen: Nach der Re-Identifikation ist kein Nachmaskieren mehr möglich. Das Nachmaskier-Werkzeug ist dann ausgeblendet, «Zurück zur Prüfung» entfällt, und ein Aufruf der Prüfseite leitet auf die re-identifizierte Fassung um. Wollen Sie etwas anderes maskieren, starten Sie einen neuen Durchgang.

10 Vorgänge verwalten

Die Seite **Vorgänge** listet alle Durchgänge der Instanz, neueste zuerst. Ein Klick auf die Bezeichnung öffnet den Vorgang wieder, inklusive Prüfseite, Nachmaskieren und Downloads. Die **Vorgang-ID** steht in der Liste und lässt sich per Klick kopieren, etwa zum Einfügen im Re-Identifikations-Formular. Alle Zeitangaben (erstellt, zuletzt geändert) erscheinen in Schweizer Lokalzeit.

SPALTE	BEDEUTUNG
Status	anonymisiert, ausgeliefert, re-identifiziert.
Integrität	Prüfung gegen das manipulationssichere Audit-Protokoll: «intakt» heisst, Original, anonymisierte Fassung und Mapping sind unverändert.
Cloud-NER-Tokens	Verbrauch der Schweizer Namenserkennungs-API für diesen Vorgang (nur gehosteter Betrieb).
Aktionen	Re-identifizieren, Mapping-CSV (mit Bestätigung), Löschen (mit Bestätigung).

Aufbewahrung: Vorgänge bleiben gespeichert, bis Sie sie löschen. Das Löschen entfernt die ganze Akte (Original, anonymisierte Fassung, Mapping) und bleibt als Ereignis im Audit-Protokoll nachweisbar.

Zugriff: Ein Vorgang ist nur für seine Erstellerin/seinen Ersteller und für Administratoren sichtbar.

11 Benutzer und Rollen (Administration)

Die Benutzerverwaltung finden Administratoren oben rechts unter **Benutzer**. Jede Instanz führt ihre eigenen Konten, ohne externen Dienst, auch offline funktionsfähig.

ROLLE	RECHTE
user	Anonymisieren, eigene Vorgänge einsehen/nachmaskieren/re-identifizieren/ löschen, Standardbegriffe pflegen, eigenes Passwort ändern.
admin	Zusätzlich: alle Vorgänge der Instanz einsehen und verwalten, Benutzer anlegen und entfernen.
• Benutzer anlegen: Benutzername + Startpasswort (mindestens 10 Zeichen) + Rolle. Maximal 5 Benutzer pro Instanz. • Passwörter werden nur als Argon2id-Hash gespeichert; das Startpasswort übermitteln Sie der neuen Benutzerin/dem neuen Benutzer auf sicherem Weg, sie/er ändert es unter Profil. • Rollen werden bei jeder Anfrage frisch geprüft: ein Rollenentzug wirkt sofort. • Der letzte verbliebene Admin lässt sich nicht löschen oder herabstufen; das Cockpit meldet das mit einem Hinweis, damit die Instanz nie ohne Administration dasteht. • Für getrennte Firmen oder Mandanten gilt: eine eigene Instanz pro Firma, nicht geteilte Konten. Sprechen Sie uns an.	

12 Sicherheit und Datenhaltung

- **Alles verschlüsselt (ruhend):** Alle Vorgangs-Artefakte (Original, anonymisierte und re-identifizierte Fassung, Quelldatei, Mapping, Metadaten) sind mit AES-256-GCM verschlüsselt und verlassen die Instanz nie; der Master-Schlüssel liegt getrennt vom Daten-Volume. Das Audit-Protokoll enthält keine Dokument-Inhalte, jedoch Benutzernamen und Vorgangs-Metadaten im Klartext.
- **Aufbewahrung:** Vorgänge bleiben bis zu Ihrer expliziten Löschung erhalten; die Löschung erfolgt vollständig und wird protokolliert.
- **Audit-Protokoll:** HMAC-verkettet und PII-frei: Umsortieren, Umschreiben oder Einfügen von Einträgen wird erkannt.
- **Egress-Kontrolle:** Eine zweite, unabhängige Prüfung vor jeder Auslieferung (siehe Kapitel 8).
- **Betrieb:** On-Premise bei Ihnen oder als isolierte Single-Tenant-Instanz in der Schweiz, betrieben durch iConference. Details: Datasheet und AVV.

13 Grenzen und Tipps

- **Kein 100%-Automat:** Die Erkennung ist mehrschichtig, aber nicht unfehlbar. Der Schutz beruht auf dem Zusammenspiel mit Ihnen: gegenlesen, nachmaskieren, Standardbegriffe pflegen.
- **Nicht unterstützt:** gescannte PDFs ohne Textebene (OCR ist auf der Roadmap) sowie Word-Dokumente mit aktiven Änderungsnachverfolgungen oder eingebetteten Objekten. Heben Sie die Nachverfolgung auf bzw. speichern Sie sauber als .docx.
- **Limiten:** 150'000 Zeichen pro Text, 25 MB pro Upload.
- **Rückweg planen:** Für Text-Antworten die .txt/.md-Fassung verwenden. Das PDF ist ein reines Versandformat. Excel-Auswertungen laden Sie als .xlsx hoch und erhalten sie zellgenau re-identifiziert zurück, Word-Dokumente als .docx strukturerhaltend.
- **Grosse Dokumente:** Die Freitext-Erkennung arbeitet in Abschnitten; bei sehr langen Dokumenten kann die Analyse einige Minuten dauern.

14 Meldungen und Hilfe

MELDUNG

BEDEUTUNG UND ABHILFE

«Free-text pass failed»	Die Freitext-Erkennung war nicht erreichbar oder hat einen Fehler gemeldet: der Vorgang bricht bewusst ab (fail-closed), es wurde nichts unsicher verarbeitet. Kurz warten und erneut versuchen; besteht das Problem, melden Sie sich beim Support.
«Gescanntes PDF»	Das PDF hat keine Textebene. Quelle als Text/Markdown bereitstellen oder vorab per OCR umwandeln.
«Format nicht unterstützt»	Erlaubt sind PDF (Textebene), Word (.docx), Markdown, Text und Excel (.xlsx).
«Upload zu gross»	Limite 25 MB: Datei aufteilen oder reduzieren.
«Nicht im Text gefunden: ...»	Der nachzumaskierende Begriff kommt im Text nicht vor. Schreibweise im Dokument prüfen (Tippfehler? Getrennt geschrieben?).
«Vorgang bereits re-identifiziert»	Abgeschlossene Vorgänge lassen sich nicht mehr nachmaskieren. Neuen Durchgang starten.
Platzhalter rot markiert	Der Platzhalter in der Antwort ist diesem Vorgang unbekannt (oft vom KI-Modell verändert). Im Original prüfen und die Stelle von Hand übernehmen.

Support: iConference AG · info@iconference.ch · iconference.ch