

ERGEBNISBERICHT

PROJEKT PASSENGER

PRIVACY ENSURING AND SECURE EXCHANGE OF BIOMETRIC RECORDS

 **Bundesministerium**
Finanzen



beauftragt im Sicherheitsforschungs-Förderprogramm KIRAS des Bundesministeriums für Finanzen

Autor: Bernhard Kohn

28.02.2025

Version 1.0

Inhaltsverzeichnis

1	Einleitung.....	3
2	Grundlagen PNR/API Daten, Verarbeitung & Tools.....	3
2.1	Allgemeiner Überblick.....	3
2.1.1	Definition und Bedeutung von PNR-Daten	3
2.1.2	Übermittelte Datenkategorien (PNR und API):	3
2.1.3	Begriffsbestimmungen im PNR-Kontext	4
2.2	Technische Details	4
2.2.1	Struktur der PNR- und API-Daten.....	4
2.2.2	Datenübermittlung und Speicherung	4
2.3	Verwendung in der EU	6
2.3.1	PNR-Daten in der Risikoanalyse	6
2.3.2	Fallbeispiele und Gerichtsurteile.....	6
2.3.3	Nutzungsbeschränkungen nach österreichischem Recht	7
2.4	Neue EU-Richtlinien und zukünftige Entwicklungen	8
2.4.1	Aktuelle EU-Vorgaben und geplante Neuerungen	8
2.4.2	Technische Anforderungen für die Umsetzung	9
2.4.3	Folgenabschätzung und Herausforderungen	9
3	Methoden und Tools.....	10
3.1	Methoden PIU.....	10
3.2	Tools Assessment	11
3.2.1	UNO/CT Go Travel	11
3.2.2	WCC/Hermes	13
3.2.3	Sita	14
3.2.4	Weitere Anbieter	15
3.2.5	Vergleich der wichtigsten Features.....	16
3.3	KI Methode Proof-of-Concept.....	17
4	Anonymes Biometrisches Matching	18
4.1	Einleitung	18
4.2	Methoden für anonymes biometrisches Matching.....	19
4.2.1	Homomorphe Verschlüsselung (HE)	19
4.2.2	Secure Multi-Party Computation (MPC)	20
4.2.3	Distributed Ledger / Blockchain	21
4.3	Umsetzung in aktuellen Forschungsarbeiten (Fallstudien)	22
4.4	Bewertung der Methoden: Eignung für anonymes biometrisches Matching	23
4.5	Fazit	25
5	Risk Assessment Toolkit und Vision	26
5.1	Einleitung	26
5.2	Proof-Of-Concept und User Interface Vision	26
6	Dissemination Aktivitäten	32
7	Literatur	33

1 EINLEITUNG

In diesem Bericht sind die Ergebnisse, die im Rahmen des Projekts Passenger erzielt wurden, dargestellt. Das Projekt Passenger ist in verschiedene 5 Arbeitspakete aufgeteilt, neben dem Arbeitspaket 1, dass das Projektmanagement beinhaltet, sind die eigentlichen Arbeiten in die Arbeitspakete AP2 – AP4 aufgeteilt.

Im Kapitel 2 werden die Grundlagen der PNR/API Datenverarbeitung und auch die Probleme, die es in der Verarbeitung bereits gegeben hat anhand von Fallbeispielen dargestellt. In Anschluss wird die generelle Arbeitsweise einer PIU aufgezeigt. Anschließend werden die verschiedenen schon existenten Lösungen beschrieben, die aus Informationen aus dem Internet und durch Demonstrationen der Anbieter gewonnen wurden. Es folgt ein kurzer Überblick weiterer Tools und als Abschluss eine Gegenüberstellung der verglichenen Systeme.

Im AP 2 sollen das Tool Assessment, KI Methoden als Proof-Of-Concept und das sogenannte Anonymous Biometric Matching untersucht werden. Das Tool Assessment und die KI Methoden werden im Kapitel 3 beschreiben. Zur besseren Gliederung ist das Anonymous Biometric Matching in einem eigenen Kapitel 4 ausgegliedert.

Im Rahmen des Risk Assessment Framework haben wir eine mögliche Erweiterung der bestehenden Methoden in der PIU ausgearbeitet und dieses wird mit Hilfe eines Chat-Bots basierend auf einem lokalen Large Language Modell (LLM) umgesetzt. Dadurch konnten wir Möglichkeiten aufzeigen, wie die PIU erhebliche Zeiteinsparungen bei der Erstellung der Kriterienabfrage erzielen kann. Von den Mitarbeitern der PIU wurde diese KI-Methode mit Begeisterung aufgenommen. Zusammen mit der User Interface Vision ist das im Kapitel 5 dargestellt.

Zum Abschluss des Berichts wird noch ein Überblick über die Dissemination Aktivitäten im Kapitel 6 gegeben. Im Literaturverzeichnis sind allfällige Links zu Webseiten oder Veröffentlichungen, aus denen Informationen gewonnen wurden, zusammengefasst.

1 GRUNDLAGEN PNR/API DATEN, VERARBEITUNG & TOOLS

1.1 Allgemeiner Überblick

1.1.1 Definition und Bedeutung von PNR-Daten

Passenger Name Record (PNR) bezeichnet einen Fluggastdatensatz, den Fluggesellschaften in ihren Buchungssystemen für jeden Passagier erstellen. Darin sind alle wesentlichen Informationen zu einer Reise zusammengefasst. PNR-Daten umfassen personenbezogene Angaben der Fluggäste, die bei der Flugbuchung erhoben werden, etwa Name, Reisedaten, Flugroute, Sitzplatz, Gepäckinformationen, Kontaktangaben (Adresse, Telefonnummer, E-Mail) und Zahlungsart [1], [2]. Diese Informationen entstehen ursprünglich im kommerziellen Kontext der Reisebuchung, werden inzwischen jedoch auch von Sicherheitsbehörden genutzt. Die Bedeutung der PNR-Daten liegt heute vor allem in ihrer Verwendung zur Gefahrenabwehr: Durch Analyse dieser Fluggastdatensätze sollen terroristische Straftaten und schwere Kriminalität verhindert, aufgedeckt und verfolgt werden [1]. PNR-Daten ermöglichen es Behörden, potenzielle *Risiko-Passagiere* zu identifizieren – also nicht nur bereits bekannte Gefährder, sondern auch bislang unbekannte Personen, deren Reiserouten oder Verhaltensmuster auf schwere kriminelle Aktivitäten hindeuten könnten [3].

1.1.2 Übermittelte Datenkategorien (PNR und API):

Die konkret übermittelten Fluggastdaten lassen sich in PNR-Daten und API-Daten unterteilen. PNR-Datensätze enthalten, wie beschrieben, umfangreiche Buchungs- und Reiseinformationen (bis zu rund 60 Datenpunkte können darin enthalten sein [4]). Dazu zählen z.B. der Buchungscode, Datum von Buchung/Ticket-Ausstellung, vollständige Reiseverläufe (inkl. Umstiege), Vielflieger-Nummern, sowie Freitextfelder mit vom Reisebüro oder Passagier angegebenen Sonderwünschen [2]. Selbst scheinbar triviale Angaben wie Mahlzeitwünsche (vegetarisch, kosher etc.) oder mitreisende Personen können Teil des PNR sein und Rückschlüsse auf persönliche Vorlieben oder Beziehungen zulassen [4]. Die Auswertung derartiger Daten (z.B. Verpflegung die kosher ist) ist laut dem österreichischen Gesetz verboten, auch wenn diese Daten vorhanden sind. Demgegenüber stehen Advance Passenger Information (API)-Daten, die als „erweiterte Fluggastdaten“ vor dem Flug erfasst und übermittelt werden. API-Daten betreffen primär die Identitätsdetails aus den Reisedokumenten der Passagiere: Name, Geburtsdatum, Geschlecht, Nationalität,

Passnummer, Ausstellungsland und Gültigkeit des Ausweises sowie reisebezogene Angaben wie Flugnummer, Abflug- und Ankunftszeitpunkt [2]. Diese Informationen werden in der Regel beim Check-in aus dem maschinenlesbaren Bereich des Reisepasses erfasst und mit Flugdaten kombiniert [1]. API-Daten gelten als verifizierte Daten, da sie direkt aus amtlichen Ausweisdokumenten stammen, während PNR-Daten von den Reisenden oder den Fluggesellschaften selbst angegeben und nicht unbedingt überprüft werden [4]. Wichtig ist, dass API-Daten oft als Teilmenge in den PNR-Datensätzen enthalten sein können – die Kategorien überschneiden sich also teilweise [1]. Beide Datentypen ergänzen einander: PNR liefert kontextreiche Buchungsinformationen, API liefert behördlich verifizierte Personendaten.

1.1.3 Begriffsbestimmungen im PNR-Kontext

Im Zusammenhang mit Fluggastdatensätzen sind einige Schlüsselbegriffe zu beachten. Fluggastdatensatz (PNR) bezeichnet wie erwähnt den digitalen Datensatz zu einer Flugreise, der in der Regel pro Buchungsvorgang angelegt wird (oft auch für Gruppenbuchungen gemeinsam). *API-Daten* sind die vorab übermittelten Passagierinformationen aus Ausweisdokumenten, die beim Boarding durch Auslesen der Datenzeilen im Reisepass erhoben werden. Die Verarbeitung dieser Daten in der EU erfolgt durch spezielle nationale Stellen, den Passenger Information Units (PIU) oder Fluggastdatenzentralstellen. Jede EU-Mitgliedstaat hat gemäß PNR-Richtlinie eine solche PIU eingerichtet [5] – in Österreich etwa ist sie beim Bundeskriminalamt angesiedelt. Die PIUs empfangen die PNR- (und API-) Daten der Fluggäste und sind für deren Speicherung, Auswertung und Weitergabe an zuständige Strafverfolgungsbehörden verantwortlich. Weiterhin unterscheidet man bei Flügen zwischen „Drittstaatsflügen“ (Flügen von oder in Drittländer, also Länder außerhalb der EU) und EU-internen Flügen. Die EU-PNR-Richtlinie 2016/681 verpflichtete die Erhebung von PNR-Daten zunächst nur für Flüge von außerhalb in die EU bzw. umgekehrt [6]. Mitgliedstaaten konnten jedoch freiwillig auch innereuropäische Flüge einbeziehen [6]. Der Begriff PNR-Richtlinie steht für jene EU-weite Regelung aus 2016, welche die Nutzung von Fluggastdatensätzen zu Sicherheitszwecken harmonisiert; sie definiert auch Schlüsselbegriffe und datenschutzrechtliche Vorgaben (etwa welche Straftaten als "schwere Kriminalität" im Sinne der PNR-Datenverwendung gelten). Insgesamt bildet dieses Begriffsgerüst den Rahmen, in dem PNR-Daten im Flugverkehr erhoben, übermittelt und verwendet werden.

1.2 Technische Details

1.2.1 Struktur der PNR- und API-Daten

Die PNR-Daten werden in den Buchungs- und Check-in-Systemen der Fluggesellschaften gespeichert und typischerweise in standardisierten Formaten an die Behörden übermittelt. Internationale Standards (wie das von IATA/ICAO entwickelte *PNRGOV*-Format auf Basis von EDIFACT/XML) legen fest, wie die Datensätze elektronisch strukturiert werden, um einen einheitlichen Austausch zwischen Airlines und staatlichen Stellen zu ermöglichen. Ein PNR-Datensatz besteht aus einer Sammlung von Feldern, die verschiedene Kategorien abdecken. Dazu gehören u.a. der eindeutige PNR-Code bzw. Buchungsnummer, Angaben zur Flugstrecke (Abflug-/Zielort, Zwischenlandungen), Flugdaten (Datum, Uhrzeit, Flugnummer), Personendaten (Name, Titel, ggf. Geburtsdatum), Kontaktdaten (Telefon, E-Mail, Anschrift), Reisedokumentdaten (Passnummer etc., falls im Buchungssystem erfasst), Zahlungsinformationen (Kreditkartennummer, Rechnungsadresse, Zahlungsmethode), Mitreisende Personen auf derselben Buchung, Sitzplatznummer, Gepäckangaben (Anzahl/ Gewicht der aufgegebenen Gepäckstücke) und ggf. Service- und Sonderwunschangaben (etwa Mahlzeitenpräferenzen, benötigte Hilfsdienste) [1] [7]. Viele dieser Felder werden von den Passagieren bei der Buchung selbst angegeben und vom System unverändert übernommen. API-Daten dagegen haben eine klar definierte Struktur, meist gemäß dem UN/EDIFACT-PAXLST-Standard: Sie enthalten pro Passagier Datenelemente aus dem maschinenlesbaren Teil des Reisedokuments (Name, Geburtsdatum, Staatsangehörigkeit, Geschlecht, Dokumentenart und -nummer, Ausstellungs- und Ablaufdatum des Dokuments) sowie Flugdaten (Start-/Zielflughafen, Flugnummer, Abflug-/Ankunftszeit). Diese API-Datensätze sind typischerweise weniger umfangreich als PNR-Datensätze, aber inhaltlich standardisierter und direkt an Identitätsdokumente geknüpft [4]. Technisch werden API-Daten heutzutage oft automatisiert erfasst, z.B. durch das Scannen des Reisepasses beim Check-in, wodurch maschinenlesbare Passinformationen digital ausgelesen und ins System übernommen werden [4]. Zusammengefasst: PNR-Daten sind heterogenere Datensätze aus Buchungssystemen mit teils freitextlichen Inhalten, während API-Daten einen festgelegten, überprüften Kern an Identitäts- und Reisedaten darstellen. Beide Datentypen werden elektronisch zusammengeführt, damit die Behörden ein möglichst vollständiges Bild jedes Flugreisenden erhalten.

1.2.2 Datenübermittlung und Speicherung

Die Übermittlung der Fluggastdatensätze von den Airlines an die staatlichen Stellen erfolgt ausschließlich mittels Push-Verfahren [8]. Das heißt, die Fluggesellschaften senden die Daten aktiv an die jeweilige nationale Fluggastdatenzentralstelle – im Gegensatz zur Pull-Methode, bei der Behörden selbst auf die Airline-

Datenbanken zugreifen würden (letzteres ist in der EU aus Datenschutzgründen unzulässig). Gesetzlich sind genaue Zeitpunkte für die Datenübermittlung festgelegt: 24 bis 48 Stunden vor der planmäßigen Abflugzeit eines Fluges müssen die PNR-Daten erstmals übermittelt werden, sowie unmittelbar nach Abschluss des Boardings (Abfertigungsschluss) ein zweites Mal [8]. Beim zweiten Transfer dürfen die Airlines auch nur die Änderungen gegenüber dem ersten Datensatz übermitteln (statt nochmals alle Daten) [8]. Diese zweistufige Übermittlung stellt sicher, dass Behörden vorab grundlegende Passagierinformationen erhalten und kurz vor Abflug alle aktualisierten Details (etwa Last-Minute Buchungen oder Änderungen bei Sitzplatz, Gepäck, No-Shows) vorliegen. In Ausnahmefällen – etwa bei aufkommenden akuten Bedrohungslagen – kann die PNR-Zentralstelle einer dringenden Anfrage zufolge auch außerhalb dieser Intervalle Datentransfers anfordern [8]. Technisch erfolgen die Übermittlungen auf sicherem elektronischem Wege nach gemeinsamen Protokollen und Datenformaten, die EU-weit vorgegeben sind [8]. Die empfangenden PIU-Systeme speichern die PNR-Daten sodann in nationalen Datenbanken zur weiteren Verarbeitung.

Die Speicherung der PNR-Daten unterliegt rechtlich geregelten Fristen und Zugriffsbeschränkungen. Gemäß EU-Vorgaben werden Fluggastdaten bis zu fünf Jahre lang aufbewahrt, allerdings mit abgestufter Zugriffsmöglichkeit [9]. In den ersten sechs Monaten nach Datenerhalt stehen alle Details im Klartext zur Verfügung, um zeitnah Ermittlungen zu ermöglichen. Danach müssen bestimmte personenbezogene Merkmale "maskiert" werden – das heißt, direkt identifizierende Felder wie Name, Kontakt- und Dokumentendaten werden durch Platzhalter oder unvollständige Angaben ersetzt [9]. Die Daten bleiben zwar gespeichert, sind aber nur noch in pseudonymisierter Form zugänglich. Eine Entmaskierung einzelner Datensätze ist dann ausschließlich unter strengen Voraussetzungen erlaubt, etwa wenn im Rahmen einer konkreten Ermittlung eine Verbindung zu diesen Flugdaten festgestellt wurde und ein autorisierter Behördenentscheider (in vielen Ländern der Leiter der PIU) die Freigabe genehmigt. Nach Ablauf der Gesamtspeicherfrist sind die PNR-Daten dauerhaft zu löschen, sofern sie nicht bereits im Zuge konkreter Fälle in Fallakten überführt wurden. (In Österreich sieht das PNR-Gesetz etwa vor, dass nach Zeitablauf *alle* Fluggastdaten endgültig zu vernichten sind, außer sie wurden für laufende Verfahren benötigt [9].) Zusätzlich zur nationalen Speicherung besteht ein geregelter Datenaustausch: relevante Treffer oder Analysen aus den PNR-Daten können zwischen den Mitgliedstaaten und mit Europol geteilt werden, um grenzüberschreitende Kriminalität effektiv zu bekämpfen [3], [6]. Dies geschieht über gesicherte Kanäle zwischen den PIUs der EU-Länder. Generell sind die Datenübermittlung und -speicherung darauf ausgelegt, große Datenmengen (im Fall Deutschlands z.B. mehrere hundert Millionen Datensätze pro Jahr [9]) automatisiert, sicher und fristgerecht zu handhaben, ohne den regulären Flugbetrieb zu stören.

Sicherheit und Datenschutzmaßnahmen: Da PNR-Daten hochsensibel sind, werden umfangreiche Schutzvorkehrungen getroffen, um Privatsphäre und Datensicherheit zu gewährleisten. Bereits bei der Datenerhebung gelten Grundsätze der Datenminimierung und Zweckbindung. Airlines dürfen nicht mehr Informationen sammeln als für die Geschäftsabwicklung nötig sind, und die Behörden dürfen nur die gesetzlich definierten Datenkategorien erheben [7]. Besonders sensible personenbezogene Daten – etwa Angaben über ethnische Herkunft, politische Meinungen, religiöse Überzeugungen, Gesundheitsdaten oder ähnliche vertrauliche Informationen – dürfen von den Behörden nicht aus PNR-Datensätzen herausgefiltert oder gespeichert werden [7]. Sollte ein Fluggast etwa spezielle Essenswünsche aus religiösen Gründen angegeben haben, dürfen solche Hinweise nicht für die Gefahrenanalyse genutzt werden, um Diskriminierung auszuschließen. Die technischen Systeme der PIUs sind darauf ausgelegt, derartige Datenfelder gar nicht erst zu verarbeiten bzw. sofort zu löschen.

Bei der Übermittlung der Daten kommen Verschlüsselungstechniken zum Einsatz, um Abhörsicherheit zu gewährleisten. Zugriff auf die PNR-Datenbanken haben nur befugte Analytiker der Fluggastdatenzentralstellen und – im Trefferfall oder auf begründete Anfrage – bestimmte Strafverfolgungsbehörden. Jeder Zugriff wird protokolliert und unterliegt der Kontrolle durch Datenschutzbeauftragte. Die Datenschutzaufsichtsbehörden der Mitgliedstaaten überwachen die Einhaltung der rechtlichen Vorgaben. In der EU-PNR-Richtlinie ist ausdrücklich festgelegt, dass die Verarbeitung der Daten „unter Einhaltung strikter Garantien für den Schutz des Privatlebens“ zu erfolgen hat [7]. Dazu zählen die erwähnten Zugriffsbeschränkungen, Pseudonymisierungsverfahren, begrenzte Speicherfristen und ein striktes Zweckbindungsprinzip: PNR-Daten dürfen ausschließlich zur Prävention, Aufdeckung, Untersuchung und Verfolgung von Terrorismus und den in der Richtlinie definierten Formen schwerer Kriminalität verwendet werden [6]. Jede darüberhinausgehende Nutzung – etwa für allgemeine Überwachungszwecke oder zur Verfolgung kleinerer Delikte – ist unzulässig. Nationales Recht (z.B. das österreichische PNR-Gesetz) präzisiert diese Grenzen und sieht teils weitere Sicherungen vor, etwa regelmäßige Evaluierungen der PNR-Stellen, Schulung des Personals im Datenschutz sowie Informationsrechte für Betroffene. Insgesamt soll durch diese technischen und organisatorischen Maßnahmen ein hoher datenschutzrechtlicher Standard sichergestellt werden [5], trotz der breiten Erfassung von Fluggastdaten.

1.3 Verwendung in der EU

1.3.1 PNR-Daten in der Risikoanalyse

In den EU-Mitgliedstaaten werden die PNR-Daten vor allem von den Passenger Information Units genutzt, um Risikoanalysen von Reisenden durchzuführen. Dabei kommen zwei Hauptansätze zum Tragen: abgleichbasierte Fahndung und musterbasierte Analyse. Zum einen werden die Fluggastdaten automatisiert mit bestehenden polizeilichen Datenbanken und Watchlists verglichen [9]. So lässt sich feststellen, ob ein Passagier zur Fahndung ausgeschrieben ist (etwa aufgrund eines Haftbefehls) oder ob seine Dokumente als gestohlen gemeldet wurden. Zum anderen werden PNR-Daten mit Hilfe von definierten Risikoprofilen durchsucht – also Kriterien- oder Regelsets, die verdächtige Reiseschemata erkennen sollen. Mit dieser Fahndungsmethode wird versucht z.B. Personen zu identifizieren, die bestimmte Auffälligkeiten in ihren Buchungen aufweisen: etwa Last-Minute-Buchungen einer Einzelflugstrecke ohne Rückflug, Bezahlung in bar oder mit vielen Zwischenstopps auf bekannten Schmuggelrouten. Trifft ein PNR-Datensatz auf genügend solcher Kriterien zu, generiert das System einen Mustertreffer. Die nationalen PIUs prüfen solche automatischen Treffer anschließend manuell nach, um Fehlalarme auszusortieren.

In der Praxis hat sich gezeigt, dass viele anfängliche Treffer keine Relevanz für die Strafverfolgung haben. Zum Beispiel wurden in Deutschland im Jahr 2022 rund 449.000 automatische Verdachtsmeldungen erzeugt (441.608 Datenbanktreffer plus 7.446 Mustertreffer) [9], doch nach eingehender Analyse blieben nur etwa 88.000 substanzielle Fälle übrig, die tatsächlich an Ermittlungsbehörden zur weiteren Verfolgung übergeben wurden [9]. Dies entspricht etwa 20% der Roh-Treffer. In ca. 19.800 Fällen konnte die Bundespolizei oder der Zoll die gesuchten Personen tatsächlich anlässlich der Einreise kontrollieren; darunter waren rund 1.387 Personen, die mit einem gültigen Haftbefehl zur Festnahme ausgeschrieben waren [9]. Diese Zahlen illustrieren den Mehrwert der PNR-Daten ebenso wie die Herausforderung: Einerseits gelingt es, mit Hilfe der Fluggastdatenauswertung eine große Zahl unbekannter Treffer herauszufiltern – darunter etliche gesuchte Straftäter, die ohne PNR-Abgleich unerkannt geblieben wären. Andererseits produziert die massenhafte Verarbeitung auch viele Fehlalarme, die personelle Ressourcen binden und datenschutzrechtlich problematisch sind. Trotzdem betonen Sicherheitsbehörden den Nutzen: Durch PNR-Analysen konnten in Europa bereits Terrorverdächtige identifiziert und Drogenschmuggler-Netzwerke aufgedeckt werden, indem z.B. Reisebewegungen auffällig wurden, die zu bekannten kriminellen Mustern passten. So ermöglicht PNR gezielte Eingriffe (etwa dass verdächtige Reisende bei Ankunft intensiver kontrolliert oder observiert werden), bevor ein Verbrechen begangen wird [7]. Die Wirksamkeit hängt jedoch stark von der Qualität der Kriterien und der internationalen Zusammenarbeit ab, da Terrorismus und Organisierte Kriminalität oft grenzüberschreitend agieren.

1.3.2 Fallbeispiele und Gerichtsurteile

Die umfassende Nutzung von Fluggastdaten hat in den letzten Jahren auch die Gerichte beschäftigt. Ein prominentes Beispiel ist die Entscheidung des Europäischen Gerichtshofs (EuGH) vom 21. Juni 2022 (Rechtssache C-817/19), welche die PNR-Richtlinie auf ihre Vereinbarkeit mit den Grundrechten prüfte [10]. Der EuGH bestätigte zwar grundsätzlich die Zulässigkeit der PNR-Datenverarbeitung zur Terrorismus- und Schwerekriminalitätsbekämpfung, forderte aber eine engere Auslegung auf das „absolut Notwendige“ [10]. Insbesondere rügten die Luxemburger Richter die bis dahin teils ausufernde Praxis der Mitgliedstaaten: So erklärten sie die pauschale Erfassung von Fluggastdaten auf innereuropäischen Flügen für unvereinbar mit EU-Recht, wenn keine „reale, aktuelle oder vorhersehbare terroristische Bedrohung“ für den betreffenden Staat besteht [10]. Mit anderen Worten dürfen PNR-Daten auf Flügen innerhalb der EU nur noch bei Vorliegen spezifischer Gefahren eingesetzt werden, nicht flächendeckend im Normalbetrieb. Weiter stellte der EuGH klar, dass die Speicherung der Daten zeitlich stärker begrenzt werden muss als in der Richtlinie vorgesehen. Eine pauschale Vorratsspeicherung aller Passagierdaten für bis zu fünf Jahre sei unverhältnismäßig, sofern die allermeisten Betroffenen keinerlei Bezug zu Straftaten haben [9], [11]. Die Richter verlangen daher etwa, dass Daten unbescholtener Reisender nach relativ kurzer Zeit (etwa 6 Monaten) gelöscht oder anonymisiert werden, und nur bei *Treffer-Personen* länger gespeichert bleiben dürfen [9]. Diese Grundsatzentscheidung hat unmittelbare Auswirkungen auf die nationalen Umsetzungen der PNR-Richtlinie.

In Deutschland wurde infolge des EuGH-Urteils z.B. die Praxis der Fluggastdatenverarbeitung angepasst. Das Verwaltungsgericht Wiesbaden urteilte 2022, dass die bisherige "Himmelsrasterfahndung" des Bundeskriminalamts – also die verdachtsunabhängige Massenauswertung aller Flugpassagierdaten – in der bisherigen Form rechtswidrig ist [2]. Daraufhin kündigte die Bundesregierung an, das deutsche Fluggastdatengesetz zu überarbeiten. Übergangsweise wurde verfügt, dass PNR-Datensätze ab 2023 in der Regel nur noch sechs Monate gespeichert werden und anschließend gelöscht oder stark eingeschränkt werden, sofern kein "verifizierter Treffer" (konkreter Fahndungshinweis) vorliegt [9]. Auch werden seitdem für EU-Binnengrenzflüge Daten nur noch bei Vorliegen einer entsprechenden Gefahreinschätzung verarbeitet,

anstatt wie zuvor routinemäßig [9]. In Österreich zeichnete sich ein ähnliches Bild ab: Die Datenschutz-NGO epicenter.works kritisierte bereits 2018, die Umsetzung der PNR-Richtlinie bedeute eine "anlasslose Vorratsdatenspeicherung" aller Flugreisenden und stelle einen Generalverdacht dar [12]. Zwar gibt es bislang kein höchstgerichtliches Urteil in Österreich, das die Fluggastdatenverarbeitung direkt stoppt, doch auch hier muss die Anwendung des PNR-Gesetzes an die strengen Vorgaben des EuGH angepasst werden. Insbesondere die anlasslose Erfassung von Inlands- und EU-Flügen dürfte eingeschränkt werden, sofern keine konkrete Terrorgefahr besteht, analog zur Luxemburger Entscheidung. Zudem sorgt die österreichische Datenschutzbehörde für die Kontrolle, dass Speicherfristen und Zugriffe im Rahmen bleiben.

Auf internationaler Ebene gab es bereits 2017 ein wichtiges juristisches Signal: Damals kippte der EuGH ein geplantes PNR-Abkommen zwischen der EU und Kanada, da es nicht ausreichend Garantien für den Schutz der Privatsphäre enthielt [12]. Dieses Urteil (Gutachten 1/15) unterstrich schon vor Inkrafttreten der EU-internen PNR-Regelung, dass Datenübermittlungen in Drittländer strengen Grundrechtsprüfungen standhalten müssen. Entsprechend sorgfältig werden etwa bestehende PNR-Abkommen mit den USA oder Australien beobachtet. Insgesamt zeigen diese Fallbeispiele und Urteile, dass zwar die sicherheitspolitische Relevanz der Fluggastdaten anerkannt wird, jedoch immer wieder Nachbesserungen zugunsten des Datenschutzes eingefordert werden. Die Gerichte betonen die Notwendigkeit eng umrissener Zwecke, kurzer Speicherfristen und wirksamer Aufsicht, um die Balance zwischen Sicherheitsinteressen und Grundrechten zu halten [10], [11].

1.3.3 Nutzungsbeschränkungen nach österreichischem Recht

Das österreichische PNR-Gesetz (Bundesgesetz über die Verarbeitung von Fluggastdaten, in Kraft seit 16. August 2018) setzt die EU-Richtlinie in nationales Recht um und enthält explizite Einschränkungen für die Verwendung der PNR-Daten. Zentrale Punkte sind dabei:

- **Zweckbindung und Straftatenkatalog:** PNR-Daten dürfen ausschließlich zur Vorbeugung, Aufklärung und Verfolgung von terroristischen Straftaten und bestimmten schweren Straftaten verarbeitet werden [6]. Welche Straftaten im Einzelnen darunterfallen, ist in einem Anhang zum PNR-Gesetz aufgelistet – im Wesentlichen handelt es sich um Delikte mit einer angedrohten Freiheitsstrafe von mindestens drei Jahren (z.B. Terrorismus, schwere Gewalt- und Sexualdelikte, Menschenhandel, Drogenhandel etc.) [3]. Die Verwendung der Daten zu anderen Zwecken (etwa allgemeinen Überwachungsmaßnahmen oder Verfolgung von Bagatelldelikten) ist unzulässig.
- **Zuständige Behörden und Zugriffsbeschränkung:** In Österreich wurde eine nationale Fluggastdatenzentralstelle (PIU) beim Bundeskriminalamt (BKA) eingerichtet, die federführend für die Datenverarbeitung zuständig ist [5]. Nur speziell befugte Beamte dieser PIU dürfen die Roh-PNR-Daten auswerten. Erkenntnisse (Treffer) dürfen dann im Bedarfsfall an die im Gesetz benannten Stellen weitergeleitet werden – das sind insbesondere die Strafverfolgungsbehörden (Kriminalpolizei, Staatsanwaltschaften, Gerichte) sowie die Zollbehörden und unter bestimmten Umständen die militärischen Nachrichtendienste (Abwehramt und Heeresnachrichtenamt) innerhalb ihres gesetzlichen Aufgabenbereichs [3]. Jede Weitergabe und jeder Zugriff ist streng zu dokumentieren.
- **Datenschutz und Löschfristen:** Das österreichische PNR-G schreibt vor, dass die Datenqualität und -sicherheit gewährleistet sein muss und die Datenschutzgrundsätze eingehalten werden [5]. Wie in der EU-Richtlinie vorgesehen, werden die Datensätze nach 6 Monaten ab dem Flug anonymisiert (maskiert) und dürfen grundsätzlich nur bis max. **5 Jahre** aufbewahrt werden. Eine längere Speicherung einzelner Datensätze ist nur zulässig, wenn sie in Zusammenhang mit laufenden Ermittlungen stehen. Zudem unterliegt die PNR-Zentralstelle der Kontrolle durch das österreichische Datenschutzreferat und den Datenschutzrat. Betroffene Reisende haben – zumindest theoretisch – Rechte nach DSGVO, etwa Auskunftsanspruch, ob zu ihrer Person PNR-Daten verarbeitet wurden, soweit dies die laufenden Ermittlungen nicht gefährdet.
- **Geografischer Anwendungsbereich:** Österreich hat die PNR-Datenübermittlung für alle Flüge von außerhalb des EU/EWR-Raums von oder nach Österreich verpflichtend gemacht (Drittstaatsflüge). Darüber hinaus hat Österreich – wie viele EU-Länder – auch innereuropäische Flüge einbezogen, indem es der EU-Kommission entsprechend Mitteilung machte (Art. 2 der Richtlinie) [6]. Somit wurden bis 2022 auch Passagierdaten z.B. auf Wien-Frankfurt oder Innsbruck-London erfasst. Nach dem EuGH-Urteil 2022 muss diese Praxis aber angepasst werden: Künftig dürften PNR-Daten auf Flügen innerhalb der EU nur noch bei Vorliegen einer konkreten Terrorgefahr verarbeitet werden. Hier bleibt abzuwarten, ob der Gesetzgeber das PNR-G entsprechend ändert oder zumindest intern per Erlass einschränkt. Bis dahin wird die PIU des BKA aus Eigeninitiative wohl einen risikobasierten Ansatz verfolgen, um dem EuGH zu genügen (ähnlich wie in Deutschland angekündigt wurde [9]).

Zusammenfassend stellen die österreichischen Vorschriften sicher, dass PNR-Daten nicht schrankenlos verwendet werden dürfen, sondern nur unter engen gesetzlichen Bedingungen. Die Etablierung einer spezialisierten Zentralstelle, die klare Zweckbindung an schwere Delikte und die Einbindung datenschutzrechtlicher Aufsicht sollen garantieren, dass die Fluggastdaten zwar für Sicherheitsbelange genutzt werden, aber die Grundrechte der Reisenden gewahrt bleiben. Dennoch bleibt – wie in der gesamten EU – die Herausforderung bestehen, diese Balance laufend zu überprüfen und ggf. nachzuschärfen, insbesondere im Lichte höchstgerichtlicher Entscheidungen.

1.4 Neue EU-Richtlinien und zukünftige Entwicklungen

1.4.1 Aktuelle EU-Vorgaben und geplante Neuerungen

Die Europäische Union arbeitet fortlaufend daran, die Nutzung von Fluggastdaten zu optimieren und gleichzeitig unionsweit einheitliche Standards zu schaffen. Im Dezember 2022 legte die EU-Kommission zwei neue Verordnungsvorschläge vor, um Lücken der bisherigen Regelungen zu schließen [1]. Diese wurden inzwischen politisch gebilligt (Trilog-Einigung Ende 2024) und dürften 2025 in Kraft treten. Im Unterschied zur PNR-Richtlinie von 2016, die von den Mitgliedstaaten in nationales Recht umgesetzt werden musste, handelt es sich hierbei um EU-Verordnungen, die unmittelbar gelten [4]. Kernziel ist es, API-Daten und PNR-Daten besser zu verzahnen und den Anwendungsbereich der Fluggastdatennutzung zu erweitern, um Terrorismusbekämpfung und Grenzsicherheit zu stärken [2]. Konkret besteht das Reformpaket aus zwei Teilen:

- Eine Verordnung zur Sammlung und Übermittlung von API-Daten für Grenzkontrollen, welche die veraltete API-Richtlinie von 2004 ablösen wird. Sie schreibt vor, dass Fluggesellschaften verbindlich eine festgelegte Liste von API-Datenelementen bei allen Fluggästen erheben und an die Grenzbehörden übermitteln [1]. Neu ist hierbei, dass wirklich *alle* Flüge erfasst werden sollen – auch Flüge innerhalb des Schengen-Raums, Charter- und Privatflüge [1]. Damit sollen bislang bestehende blinde Flecken beseitigt werden. Zudem wird betont, dass die Erhebung automatisiert erfolgen muss: Durch Scan des elektronischen Reisepasses (MRZ/Chip) sollen *maschinenlesbare* Passagierdaten gewonnen werden [4], was die Qualität und Vollständigkeit der Daten erhöht. Diese API-Verordnung zielt in erster Linie darauf ab, Grenzübertrittskontrollen effizienter zu machen und illegale Einreisen frühzeitig zu erkennen.
- Eine Verordnung über die Nutzung von API-Daten zu Strafverfolgungszwecken, welche eng mit ersterer verzahnt ist [1]. Sie ermöglicht es, die erhobenen API-Daten auch für Risikoanalysen und Polizeizwecke einzusetzen – im Prinzip eine Erweiterung des PNR-Systems. Die *Basis-Identitätsdaten* der Reisenden aus den Ausweisdokumenten sollen dabei mit den umfangreicheren PNR-Buchungsdaten kombiniert ausgewertet werden [2]. So entsteht ein vollständiger Datensatz pro Passagier, der sowohl verifizierte Stammdaten (API) als auch Kontextinformationen (PNR) enthält. Durch diese Verknüpfung erhofft man sich präzisere Treffer bei der Mustererkennung. Die zweite Verordnung überträgt im Grunde die PNR-Methodik (Gefahrenabwehr, Datenabgleich mit Kriterien) auch auf die neu strukturierten API-Daten. Sie enthält Vorgaben, *wie* die Daten zu Verhütung, Aufdeckung und Verfolgung schwerer Straftaten genutzt werden dürfen, und schafft damit erstmals einen EU-weit einheitlichen Rechtsrahmen für den Einsatz von API-Daten in der Strafverfolgung – etwas, das bisher fehlte [1].

Ein zentrales Element der neuen EU-Vorgaben ist die technische Infrastruktur für den Datenaustausch: Geplant ist die Einrichtung eines zentralen „Router“-Systems, betrieben von der EU-Agentur eu-LISA, über das künftig sämtliche Fluggastdaten laufen sollen [4] [2]. Dieser Router dient als einziger Zugangspunkt, an den die Airlines ihre PNR/API-Daten senden. Von dort werden die Daten automatisiert an die zuständigen nationalen Stellen weiterverteilt. Der Vorteil besteht darin, dass Fluggesellschaften nicht mehr individuelle Verbindungen zu Dutzenden PIUs in ganz Europa managen müssen, sondern nur noch zu einer zentralen EU-Schnittstelle – was den Verwaltungsaufwand und technische Komplexität deutlich reduziert [2]. Zudem verspricht man sich dadurch einen konsistenteren Datenschutz und Datensicherheit, da der Router als "single point of entry" besser überwacht und gehärtet werden kann. Die Verordnungen legen auch präzise Zeitpunkte für künftige Datenübermittlungen fest: Vorgesehen sind zwei obligatorische Transfers, einmal beim Check-in (vor Boarding) und ein weiteres Mal nach Abschluss des Boardings [1]. Diese Klarstellung harmonisiert die bisher unterschiedlichen nationalen Praktiken (einige Länder hatten 24h vor Abflug und kurz vor Abflug, andere mehrmals gestaffelt). Insgesamt führen die neuen EU-Vorgaben dazu, dass alle relevanten Flugreisendendaten – von der Passnummer bis zur Gepäckanzahl – standardisiert, vollständig und frühzeitig den Behörden vorliegen.

1.4.2 Technische Anforderungen für die Umsetzung

Die geplanten Neuerungen stellen Airlines und Behörden vor erhebliche technische Aufgaben. Fluggesellschaften müssen ihre IT-Systeme entsprechend aufrüsten. Insbesondere muss jede Airline ein System zur automatisierten Erfassung der Ausweisdaten aller Passagiere implementieren (sofern nicht bereits vorhanden). In der Praxis bedeutet dies z.B. Installation von Dokumenten-Scannern am Check-in oder Integration mobiler Scan-Lösungen für Online-Check-ins, damit die maschinenlesbaren Passdaten zuverlässig ausgelesen werden können. Weiterhin müssen die Carriers ihre Systeme an den zentralen EU-Router anschließen [2]. Das erfordert die Entwicklung kompatibler Schnittstellen (APIs im IT-Sinn) nach den von eu-LISA vorgegebenen Protokollen. Gerade für kleinere oder Nicht-EU-Fluggesellschaften dürfte dies eine Herausforderung sein, da sie bisher ggf. nur vereinfachte Verfahren genutzt haben. Eine Übergangsphase ist daher vorgesehen, in der notfalls noch manuelle Datenerfassungen zulässig sind [2] – doch mittel- bis langfristig wird die vollautomatische Echtzeit-Übermittlung zum Standard.

Auch auf Behördenseite sind Anpassungen nötig: Die nationalen PIUs müssen mit dem EU-Router kommunizieren können und ihre Datenbanksysteme so ausrichten, dass sie die angereicherten API+PNR-Datensätze verarbeiten. Das Zusammenführen von API- und PNR-Daten erfordert eventuell neue Software-Module, die z.B. anhand von Buchungsreferenzen oder Ticketnummern die Datensätze matchen. Zudem entstehen höhere Datenvolumina, da durch die Einbeziehung aller Flüge (inklusive innereuropäischer Flüge und kleinerer Airlines) deutlich mehr Datensätze anfallen werden als bisher. Die IT-Systeme müssen daher skaliert werden (Stichwort Big-Data-Management), um etwa Verzögerungen bei der Risikoanalyse zu vermeiden. eu-LISA selbst steht vor der Aufgabe, den zentralen Router technisch bereitzustellen und zu betreiben. Dies umfasst robuste Server-Infrastruktur, Cybersecurity-Maßnahmen und redundante Netzwerke, um 24/7-Betrieb zu gewährleisten – ähnlich den anderen Großsystemen, die die Agentur bereits betreibt (wie Schengen-Informationssystem, Eurodac, etc.). Schließlich müssen alle Beteiligten umfangreiche Tests und Schulungen durchführen, damit zum Startzeitpunkt der neuen Verordnungen ein reibungsloser Datenfluss gewährleistet ist. Zusammengefasst verlangen die neuen EU-Regeln erhebliche Investitionen in Technik und Know-how, sowohl seitens der Luftfahrtunternehmen als auch der staatlichen Stellen, um die ambitionierten Vorgaben in die Praxis umzusetzen.

1.4.3 Folgenabschätzung und Herausforderungen

Die Ausweitung und Zentralisierung der Fluggastdatenverarbeitung bringt Chancen, aber auch Herausforderungen mit sich. Aus sicherheitspolitischer Sicht verspricht das Reformpaket eine Effizienzsteigerung: Einheitliche Datensätze und ein zentrales Verteilersystem sollen dafür sorgen, dass relevante Informationen schneller und vollständiger bei den Ermittlern ankommen [2]. Grenzkontrollen könnten durch automatisierte Überprüfungen im Hintergrund beschleunigt werden, da Risikoreisende bereits vor Ankunft identifiziert sind. Die lückenlose Erfassung aller Flugbewegungen (inkl. innereuropäischer) soll verhindern, dass Kriminelle die bisherigen Ausnahmen (z.B. Umstieg via Schengen-Inland) ausnutzen. Eine Folgenabschätzung der EU-Kommission hat diese Vorteile betont und auf bestehende Defizite verwiesen: Bislang konnten Mitgliedstaaten selbst entscheiden, ob sie API-Daten nutzen – was zu ungleichen Vorgehensweisen und Sicherheitslücken führte [1]. Die neuen Vorgaben schließen diese Lücke durch Harmonisierung. Auch für Fluggesellschaften könnte langfristig ein einheitliches System weniger bürokratischen Aufwand bedeuten als unterschiedliche nationale Anforderungen.

Dennoch gibt es erhebliche Herausforderungen zu meistern. Datenschutz und Grundrechte bleiben ein zentrales Thema. Kritiker monieren, dass die Ausweitung auf alle Flüge faktisch eine noch umfangreichere Massenüberwachung des Reiseverkehrs bedeutet [10]. Zwar sollen die neuen Regelungen dem EuGH-Urteil von 2022 Rechnung tragen, doch die Praxis wird zeigen müssen, ob die "Notwendigkeit" jeder Datenerhebung wirklich gewahrt wird [4]. Insbesondere die dauerhafte Erfassung innereuropäischer Flüge könnte erneut vor Gericht angefochten werden, falls keine restriktiven Anwendungsfilter eingebaut werden. Die EU-Verhandler haben versucht, die Verordnungen grundrechtskonform zu gestalten – etwa durch klare Zweckbindungen und Eingrenzungen. So wurde diskutiert, PNR/API-Analysen auf bestimmte Risikomuster oder Routen zu konzentrieren, anstatt wahllos alle Daten gleich intensiv auszuwerten [9]. Dennoch bleibt abzuwarten, ob diese Maßnahmen ausreichen, um den strengen Verhältnismäßigkeitsgrundsatz des EuGH zu erfüllen.

Auch datentechnisch könnte die riesige Datenflut kontraproduktiv werden. Sicherheitsbehörden stehen vor der Aufgabe, aus Millionen zusätzlicher Datensätze die *relevanten* herauszufiltern, ohne in falschen Treffern zu ertrinken. Experten warnen, dass „zu viele Daten die Analyse schwieriger machen“ können [12] – sprich: die Signal-zu-Rauschen-Relation verschlechtert sich, wenn man noch mehr unverdächtige Fälle mitdurchsucht. Es wird also darauf ankommen, intelligente Filter und Algorithmen einzusetzen und kontinuierlich zu verbessern, um Effizienz und Trefferquoten zu erhöhen. Eine weitere Herausforderung ist die IT-Sicherheit des zentralen Routers. Dieser wird ein höchst attraktives Ziel für Cyberangriffe darstellen, da er die Daten praktisch aller Fluggäste in der EU bündelt. Ein erfolgreicher Angriff oder ein Datenleck dort hätte EU-weite

Auswirkungen. Entsprechend hoch sind die Anforderungen an die Absicherung (Firewalls, Zugriffskontrolle, Monitoring) und an die Notfallkonzepte, sollte das System ausfallen.

Nicht zuletzt stellt sich die Frage der Akzeptanz. Airlines müssen die neuen Pflichten implementieren und könnten die Kosten an die Passagiere weitergeben; Fluggäste wiederum könnten Datenschutzbedenken haben, wenn noch mehr persönliche Daten vorab erhoben werden. Transparenz gegenüber der Öffentlichkeit wird wichtig sein, um das Vertrauen in diese Maßnahmen zu erhalten. Hier spielt auch die Folgenabschätzung im rechtlichen Sinne eine Rolle: Die EU-Institutionen haben zu beurteilen, ob der erwartete Sicherheitsgewinn in einem angemessenen Verhältnis zum Eingriff in die Privatsphäre steht (Stichwort Verhältnismäßigkeitsprüfung). Diese Bewertung ist komplex und teils politisch umstritten, da Sicherheitsnutzen schwer quantifizierbar sind.

Zusammenfassend bewegen sich die zukünftigen Entwicklungen bei Fluggastdatensätzen in einem Spannungsfeld: Auf der einen Seite stehen verstärkte internationale Zusammenarbeit, modernisierte Technik und umfassendere Überwachungskapazitäten, die ein Plus an Sicherheit versprechen. Auf der anderen Seite müssen demokratische Gesellschaften sicherstellen, dass diese Maßnahmen kontrolliert, rechtsstaatlich eingehegt und effektiv sind – damit Freizügigkeit und Privatsphäre nicht unnötig eingeschränkt werden. Die EU setzt mit den neuen Verordnungen einen weiteren Schritt in Richtung „intelligenter Grenzen“ und Datenverbundsysteme (Stichwort eu-LISA Programme), was voraussichtlich auch den zukünftigen Diskurs prägen wird. Denn Fluggastdaten bleiben ein zweiseitiges Schwert: Sie können wertvolle Hinweise liefern, erfordern aber stets den Ausgleich zwischen Sicherheit und Freiheit. Die kommenden Jahre – mit der Umsetzung der neuen Vorgaben und möglichen weiteren Gerichtsurteilen – werden zeigen, wie dieser Ausgleich in der Praxis gelingt.

2 METHODEN UND TOOLS

2.1 Methoden PIU

Gestartet wurden die Arbeiten in diesem Projekt mit der Bedarfserhebung. Dazu wurde in Besprechungen erhoben, wie die Arbeitsweise in der Passenger Information Unit aussieht.

Der erste Schritt ist das Verständnis welche Daten zur Verfügung stehen. Dabei wird unterschieden zwischen den PNR und den API-Daten. PNR steht für Passenger Name Record. Diese Daten werden bei Buchung der Reise von der entsprechenden Buchungssoftware erhoben und gespeichert. Dabei werden neben den reinen Flugdaten auch sämtliche Änderungen der Buchung und auch die Daten von Begleitpersonen gespeichert. Darin enthalten sind sämtliche Daten, die bei der Buchung relevant waren, z.B. auch Daten zur Hotel- oder Mietwagenbuchung. Die PNR-Daten werden typischerweise 24 bis 48 Stunden vor dem Flug an die PIU übermittelt. Dann gibt es noch die Daten, die bei der Passagierabfertigung (Check-in) erhoben werden. Diese werden mit API bezeichnet, wobei die Abkürzung für Advanced Passenger Information steht. Die API-Daten werden meist kurz nach dem Check-in übermittelt. Zusätzlich werden sie beim Boarding, nachdem man sein Ticket vorgezeigt hat (also bei dem Scan des Barcodes) auch nochmal übermittelt.

Die Daten werden in eine Datenbank eingefügt und dabei auf Gültigkeit geprüft. Insbesondere die PNR-Daten enthalten häufiger auch falsche Daten, da wird dann das Geburtsdatum in ein falsches Datenfeld eingetragen, oder die Namen sind nicht richtig, manchmal wird an einem Namen ein „mr“ angehängen, obwohl es eigentlich nur die Anrede darstellen soll.

Die Daten werden auf eine Vielzahl von verschiedenen Fehlertypen geprüft und wenn möglich korrigiert (z.B. Geburtsdaten in ein einheitliches Format überführt). Danach werden die verwendbaren Daten in eine PIU interne Datenbank importiert.

Die tatsächlichen Hauptaufgaben der PIU lassen sich dann in drei unterschiedliche Abfragetypen unterscheiden

- Fahndungsabfrage: dabei wird mittels Namen und Geburtsdatum nach Personen in der Datenbank von Organisation wie Europol oder Interpol gesucht
- Einzelabfragen: dabei wird mittels Namen und Geburtsdatum nach besonderen Personen in der Datenbank gesucht, die nicht in den Datenbanken für die Fahndungsabfragen stehen.
- Kriterien Abfrage: dabei werden z.B. bestimmte Routen abgefragt oder z.B. inwieweit dort Personen eines bestimmten Alters nur mit Handgepäck und in Begleitung von jungen Frauen begleitet werden. D.h. es wird nach bestimmten, derzeit bekannten Auffälligkeiten, gesucht. Diese Abfragen werden mit Hilfe der Zollbehörden vom Flughafen erstellt, die diese Beobachtungen entsprechend gemacht haben.

Die Kriterien Abfrage wird momentan mit Hilfe des Tools SAP Business Objects umgesetzt, da dieses Tool in der PIU zur Verfügung steht und damit ein Mapping von Business Objects auf SQL-Datenbankentabellen ermöglicht wird. Zusätzlich kann die Fähigkeit von SAP Business Objects zur Generierung von Reports verwendet werden. Das Mapping ist allerdings recht aufwändig und benötigt ein tieferes Verständnis von SAP Business Objects und der Verwendung von Datenbanktabellen. Insgesamt wird der benötigte Zeitaufwand einer komplexeren Abfrage auf etwa 1 bis 2 Wochen Arbeitszeit geschätzt.

Die PIU selbst ist 24 Stunden 7 Tage die Woche erreichbar, so dass Fahndungsabfragen und Einzelabfragen immer abgearbeitet werden können und auch die Ergebnisse der Kriterien Abfrage entsprechend bearbeitet werden können und an die zuständigen Abteilungen weitergeleitet.

2.2 Tools Assessment

Als kommerzielle Systeme, die es ermöglichen können, eine einfachere Kriterien Abfrage umzusetzen, kommen verschiedene System in Frage. In dieser Studie konzentrieren wir uns auf die folgenden drei Systeme, die am weitesten Verwendung finden:

- Software "goTravel" Hersteller: UN-OCT
- Software "HERMES" Hersteller: WCC
- Software "Intelligence and Targeting" Hersteller: Sita

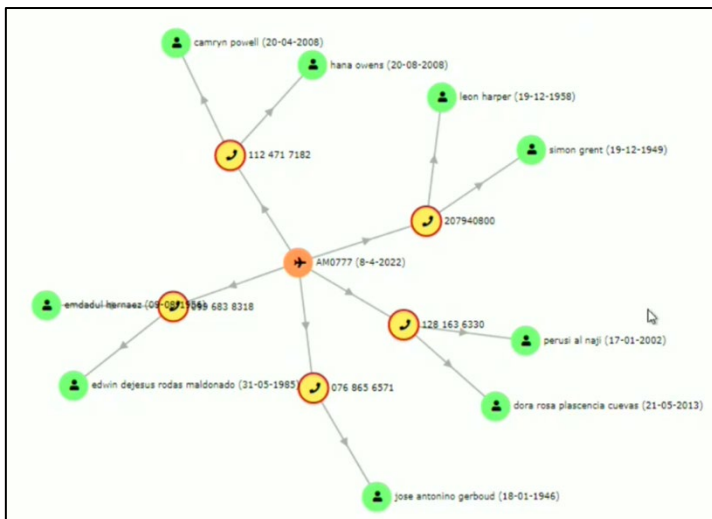
Es wurden mit allen drei Anbietern Gespräche über ihre Software geführt. Jeder der Anbieter hat uns eine Demonstration der Systeme ermöglicht und wir konnten mit Testdaten die Verwendung testen. Aufgrund der Natur der Daten können wir nicht mit einem System im Echtbetrieb arbeiten, da diese nur Mitarbeitern der PIU Einheit vorbehalten ist. D.h. wir konzentrieren uns auf die Features und Möglichkeiten der Systeme.

2.2.1 UNO/CT Go Travel

goTravel ist eine PIU-Softwarelösung, die vom United Nations Office of Counter-Terrorism (UNOCT) im Rahmen des Countering Terrorist Travel Programms bereitgestellt wird. Sie richtet sich insbesondere an UN-Mitgliedstaaten, die aufgrund internationaler Vorgaben ein eigenes API/PNR-System einführen müssen. goTravel wurde ursprünglich aus dem niederländischen *Travel Information Portal (TRIP)* entwickelt und als UN-eigenes Softwareprodukt weitergeführt [13]. Die Lösung wird Mitgliedstaaten kostengünstig zur Verfügung gestellt [14] – das heißt, UNOCT unterstützt die Implementierung (inkl. Installation, Schulung und Anbindung an Airlines), für die Wartung fallen entsprechende Wartungskosten an. Die Daten selbst verbleiben stets im Hoheitsbereich des jeweiligen Staates; die UN hat keinen Zugriff auf die laufenden Passagierdaten [14]. goTravel ermöglicht den vollständigen End-to-End-Prozess vom Dateneingang bis zur Risikoanalyse und bildet das Herzstück einer PIU, um verdächtige Reisende zu identifizieren und Informationen für Strafverfolgungsbehörden aufzubereiten [13]. Über 50 Staaten haben bereits Interesse an der Nutzung von goTravel angemeldet oder das System in Betrieb [14].

Funktionsweise: Die goTravel-Software umfasst alle wesentlichen Funktionen eines PIU-Systems. Zunächst agiert sie als zentrale Datendrehscheibe, die API- und PNR-Nachrichten in unterschiedlichen Formaten empfangen kann (z.B. PNRGOV in EDIFACT oder XML, UN/EDIFACT PaxLST) [14]. Das macht die Anbindung verschiedener Fluggesellschaften und Buchungssysteme flexibel. Eingehende Passagierdaten werden in der nationalen Datenbank gespeichert und automatisch analysiert. goTravel erlaubt den PIU-Analysten, eigene Risk Indikatoren (Regeln) zu konfigurieren und behördliche Watchlists (z.B. Terrorismusverdächtige, Haftbefehle, gestohlene Pässe) einzubinden [14]. Anhand dieser Kriterien werden alle Passagiere vor ihrer geplanten Ankunft oder Abreise risikobewertet [14]. Ein Abgleich mit Interpol-Daten (etwa gestohlenen Reisedokumenten) ist ebenfalls integriert [14]. Ergibt die Prüfung Treffer – d.h. ein Datensatz entspricht einer Regel oder Watchlist-Eintragung – so wird dies im System vermerkt. Die zuständigen Behörden können automatisch benachrichtigt werden, wenn goTravel eine Person mit hohem Risiko identifiziert (z.B. per E-Mail oder Schnittstelle) [14]. Neben der automatischen Vorab-Screening-Funktion bietet goTravel Ermittlungswerkzeuge: Benutzer können die gesammelten API/PNR-Daten manuell abfragen und durchsuchen, um etwa in laufenden Untersuchungen Informationen über Verdächtige und ihre Reisehistorie zu gewinnen [14]. Diese Query-Funktion ist für die Strafverfolgung wertvoll, um auch nachträglich Bewegungsprofile von Personen zu analysieren. Insgesamt arbeitet goTravel nach dem Prinzip "Targeting the known and the unknown" – also bekannte Gefährder durch Regeln/Listen aufzuspüren und unbekannte Verdachtsmomente durch Datenkombination sichtbar zu machen [14].

KI-gestützte Analyse: Im Gegensatz zu den anderen beschriebenen kommerziellen Lösungen ist goTravel primär regelbasiert, bietet aber dennoch einige advanced analytics. Besonders hervorzuheben ist das Network Analysis-Modul: Die Software kann Beziehungen zwischen Datenpunkten erkennen und visualisieren, um bislang verborgene Verbindungen aufzudecken [14]. Beispielsweise könnte goTravel feststellen, dass



mehrere Reisende die gleiche Telefonnummer oder Kreditkarte angegeben haben – ein möglicher Hinweis auf kriminelle Netzwerke. Solche Beziehungen werden den Analytikern über graphische Darstellungen zugänglich gemacht (siehe seitlich). Durch diese netzwerkbasierende Analyse können Muster identifiziert werden, die über einfache Regeln hinausgehen ("formally unknown relationships" [14]). Darüber hinaus ermöglicht goTravel den Austausch von Intelligence Packages – d.h. Erkenntnisdokumenten – zwischen Behörden: Auffällige Fälle können pakettiert und an Strafverfolgungsbehörden oder international geteilt werden [14]. Dieses Feature unterstützt insbesondere die Weiterverfolgung von Treffern (z.B. wenn ein

PIU einen potentiellen Terroristen identifiziert, können die Informationen gebündelt an die Polizei zur Intervention übergeben werden). goTravel ist darauf ausgelegt, Best Practices aus verschiedenen Ländern zu vereinen; durch die UNOCT-Initiative werden Erfahrungen aus mehreren Staaten gesammelt und fließen in die Weiterentwicklung des Systems ein [14]. Zwar ist KI (im Sinne selbstlernender Algorithmen) in goTravel nicht enthalten, doch das System ist offen für Erweiterungen. Durch die Zusammenarbeit der goTravel International User Community (IUC) werden künftige Features – etwa weitergehende automatisierte Mustererkennung – gemeinsam priorisiert und umgesetzt [15]. Schon jetzt erlaubt goTravel eine Kombination von Indikatoren, die es ermöglicht, neben bekannten Szenarien auch neuartige Verdachtsmomente zu erkennen (Targeting of "unknowns") [14]. Insgesamt bietet goTravel damit eine solide, wenn auch nicht unbedingte KI-Unterstützung, welche jedoch laufend ausgebaut werden kann.

Benutzerfreundlichkeit: Als von der UN betreute Plattform wird Wert daraufgelegt, dass goTravel in verschiedenen Ländern mit unterschiedlichen Ressourcen einsetzbar ist. Die Software wird auf den Systemen des Mitgliedstaats installiert und wartungsarm gehalten – UNOCT unterstützt bei Updates und Fehlerbehebung im Rahmen von Service-Vereinbarungen [14]. Die Oberfläche ist webbasiert und orientiert sich an Standard-Layouts: In einer zentralen Dashboard-Ansicht sieht der Nutzer die wichtigsten Kennzahlen und Alarmer. Die Einrichtung von Regeln und Watchlists erfolgt über konfigurierbare Masken, was den operativen Anwendern (z.B. Polizeibeamten in der PIU) ermöglicht, ohne Programmierung die Suchkriterien zu steuern [14]. Zu beachten ist, dass eine derartige in dem System angelegte Watchlist in Österreich nach dem PNR Gesetz verboten ist. Das System wurde mit Blick auf Datenschutz entwickelt – es implementiert Prinzipien wie *Privacy by Design*, sodass beispielsweise eine Trennung zwischen sensiblen und weniger sensiblen Daten vorgenommen werden kann [14]. Auch verfügt goTravel über Module zur Datenweitergabe: Innerhalb eines Landes können verschiedene Behörden (z.B. Grenzpolizei, Nachrichtendienst) über einen Berechtigungsmechanismus auf die Plattform zugreifen, zudem ist ein optionales Austausch-Modul vorgesehen, um zwischen befreundeten Staaten goTravel-Daten zu teilen (etwa im Rahmen regionaler Abkommen) [14]. Die Benutzerfreundlichkeit zeigt sich ferner in der ausführlichen Schulung: UNOCT stellt Training und Mentoring bereit, damit lokale Nutzerteams die Software eigenständig bedienen und pflegen können. Nach einer Einführungsphase können die Länder das System eigenständig weiterführen, wobei ein Nutzertreffen (IUC) den Erfahrungsaustausch erleichtert [15]. Insgesamt ist goTravel auf Pragmatismus und einfache Handhabung getrimmt – es erfüllt die Kernanforderungen, ohne überflüssigen Ballast, und lässt sich an nationale rechtliche Vorgaben anpassen [14]. Dies macht es gerade für Staaten attraktiv, die begrenzte Mittel haben, aber die internationalen PNR-Verpflichtungen erfüllen müssen.

Grafische Darstellungen und Interface: Obwohl goTravel als kostenfreie Lösung konzipiert ist, bietet sie moderne Visualisierungsmöglichkeiten. Besonders hervorzuheben ist die Grafische Netzwerk-Analyse: Analytiker können mittels eines Graphen auf dem Bildschirm schnell erfassen, welche Passagiere untereinander Verbindungen teilen (z.B. gleiche Kontaktinformationen), was auf verdeckte Zusammenhänge schließen lässt [14]. Darüber hinaus stellt goTravel tabellarische und kartografische Ansichten bereit. Verdächtige Passagiere können mit ihren Reiserouten auf einer Weltkarte angezeigt werden, um Bewegungsmuster zu erkennen. Eine Kartenübersicht der aktuellen Flüge und deren Risikoauswertung verschafft dem PIU einen Überblick über alle laufenden Reisen. Auch statistische Auswertungen (Anzahl der Treffer, verarbeitete PNRs pro Zeitraum, etc.) sind integriert, damit der Nutzen der Maßnahmen dokumentiert werden kann. Insgesamt deckt die Benutzeroberfläche die zentralen Anforderungen ab, auch wenn sie im Vergleich zu kommerziellen Produkten weniger verspielt ist. Für viele Anwender steht jedoch die Zuverlässigkeit im Vordergrund – und hier bietet goTravel solide Performance bei der Visualisierung von Risiken vor der Ankunft eines Passagiers.

2.2.2 WCC/Hermes

WCC HERMES ist eine Lösung des niederländischen Anbieters WCC und stellt ein personenzentriertes, integriertes Grenzsicherheits-System zur Überprüfung von API- und PNR-Daten dar [16]. Ziel ist es, Sicherheitsbedrohungen effektiv zu erkennen, ohne den Passagierfluss zu beeinträchtigen [16]. HERMES fungiert dabei als zentrale Plattform, die den Anschluss sämtlicher Transportträger (Fluglinien, aber auch See-, Bahn- und Busunternehmen) erleichtert und unterschiedliche Nachrichtenformate (PNRGOV EDIFACT, XML, PAXLST etc.) verarbeiten kann [14]. Nach Empfang der Passagierlisten extrahiert HERMES alle relevanten Informationen und führt automatisierte Abgleiche durch (siehe unten). Die Lösung kann sehr schnell implementiert werden (typischerweise innerhalb eines Tages) und skaliert auf große Passagieraufkommen [16].

Funktionsweise: HERMES ermöglicht den PIUs eine umfassende automatisierte Risikoprüfung aller eingehenden Passagierdaten. Zunächst werden die API/PNR-Datensätze von den Fluggesellschaften oder anderen Beförderern entgegengenommen und geparkt. Anschließend erfolgt ein Abgleich mit diversen Watchlists (z.B. polizeiliche Fahndungslisten, Interpol-Datenbanken) mittels leistungsfähiger Algorithmen für multi-kulturelle Namensabgleiche [16]. WCC's ELISE-Plattform wird genutzt, um auch bei abweichenden Schreibweisen oder unvollständigen Daten zuverlässige Treffer zu erzielen [16]. Neben solchen regelbasierten Trefferprofilen (Risk Profiles) können die Grenzbehörden eigene Zielindikatoren definieren und flexibel anpassen [16]. HERMES bewertet Passagiere bereits vor Ankunft bzw. Abflug anhand dieser Risikoprofile und Watchlist-Treffer und markiert verdächtige Personen für weitere Überprüfungen [14]. Für die Ermittlungsarbeit steht zudem eine Recherchekomponente zur Verfügung, mit der Analysten manuell in den gesammelten API/PNR-Daten suchen können (z.B. historische Reiserouten über bis zu 5 Jahre) [16] [14]. Werden potenzielle Risiken identifiziert, kann das System automatisiert die zuständigen Behörden alarmieren [14], damit am Flughafen oder Grenzübergang entsprechende Maßnahmen (wie sekundäre Kontrollen oder Einreiseverweigerung) eingeleitet werden.

KI-gestützte Analyse: WCC HERMES war eine der ersten PIU-Lösungen, die Künstliche Intelligenz in die Passagierdatenauswertung integrierte [17]. KI-Module ergänzen die klassischen Regel- und Watchlist-Ansätze, indem sie Anomalieerkennung betreiben: Ein Algorithmus analysiert alle eingehenden Passagierdatensätze und flaggt Reisende mit ungewöhnlichen oder auffälligen Merkmalen [17]. Dadurch können auch bisher unbekannte Muster oder neuartige Bedrohungen erkannt werden, die durch feste Regeln allein nicht abgedeckt sind. Zusätzlich nutzt HERMES KI für Mustererkennung und Trendanalysen im vorhandenen Datenbestand [17]. So kann etwa prognostiziert werden, wie viele Treffer (Matches) an einem bestimmten Tag zu erwarten sind, was wiederum die Personalplanung der Grenzbehörden unterstützt [17]. Wichtig ist, dass die KI dabei unterstützend wirkt: Die endgültige Entscheidung liegt stets beim Menschen, während die KI hilft, Auffälligkeiten schneller und präziser zu finden [17]. HERMES kann zudem anhand von Machine-Learning-Algorithmen Vorschläge für neue Risiko-Profile generieren (etwa wenn wiederkehrende Anomalien erkannt werden) [16]. Insgesamt verspricht die KI-Integration eine schnellere, sicherere und verlässlichere Bedrohungserkennung an der Grenze [17].

Benutzerfreundlichkeit: Die Bedienoberfläche von HERMES wurde in enger Abstimmung mit erfahrenen Grenzbeamten entwickelt, um eine möglichst intuitive Nutzung zu gewährleisten [16]. Die Software ist *out-of-the-box* einsatzbereit für API und PNR und erfordert nur minimalen Schulungsaufwand. So erlaubt die Benutzeroberfläche den Anwendern (z.B. PIU-Analysten oder Grenzbeamten), Trefferfälle leicht nachzuverfolgen, Kommentare zu hinterlegen, weitere Fahndungslisten einzupflegen oder eigene neue Regeln (Targeting-Profilen) anzulegen [16]. All dies kann ohne tiefgehende IT-Kenntnisse erfolgen, wodurch die Behörden flexibel und unabhängig bei der Konfiguration ihrer Risikoindikatoren sind. Die Oberfläche bietet One-Stop-Übersichten, in denen alle relevanten Informationen zu einem Treffer auf einen Blick ersichtlich sind [16]. Dank ausgefeilter Such- und Filtermöglichkeiten lassen sich auch große Datenbestände (Millionen von Datensätzen) schnell durchsuchen und analysieren [16]. HERMES ist webbasiert und modular aufgebaut; Updates mit neuen Funktionen werden WCC-Kunden regelmäßig bereitgestellt [16]. Für Betrieb und Wartung bietet WCC 24/7 Support, was die Lösung insbesondere für kritische Grenzkontroll-Infrastrukturen attraktiv macht [16].

Grafische Darstellung und Interface: HERMES stellt verschiedene Werkzeuge zur Datenanalyse visuell bereit. So werden zum Beispiel zur Mustererkennung Heatmaps und andere Analyse-Dashboards angeboten, um Auffälligkeiten im Reiseverkehr zu visualisieren [16]. Treffer können nach Flug, Route oder anderen Kriterien gruppiert und graphisch dargestellt werden. Ferner unterstützt das System die Visualisierung komplexer Zusammenhänge zwischen Datenpunkten durch Export an spezifische externe Tools wie z.B. den IBM i2 Analyst's Notebook (siehe kurze Erläuterung in Kapitel 3.2.4 Weitere Anbieter). Etwaige Verbindungen zwischen Passagieren – z.B. gemeinsame Adressen, Telefonnummern oder Kreditkarten – lassen sich erkennen, was bei der Aufdeckung von Netzwerken (z.B. Schleuserringen) hilft [14]. Insgesamt legt WCC erkennbar Wert auf eine benutzerzentrierte Oberfläche: Die wichtigen Informationen werden konsolidiert

präsentiert, und zusätzliche Module (z.B. für Statistik oder Berichtswesen) können integriert werden. Damit ist HERMES als flexible Komplettlösung positioniert, die von der Datenerfassung bis zur Risikoanalyse alle Schritte abdeckt [16].

2.2.3 Sita

SITA Intelligence and Targeting (häufig im Kontext von SITA auch als Teil der *iBorders* bzw. *SITA at Borders*-Lösungen bezeichnet) ist die PIU-Software des globalen IT-Dienstleisters SITA, der auf Technologien für Luftverkehr und Grenzen spezialisiert ist. SITA's Lösung ermöglicht es Regierungen, Risiken an der Grenze in Echtzeit zu erkennen und zu bewältigen [18]. Sie wurde entwickelt, um große Mengen an Reise- und Passagierdaten in verwertbare Erkenntnisse („actionable intelligence“) umzuwandeln [18]. SITA Intelligence and Targeting integriert API- und PNR-Daten aller Verkehrswege (Luft, Land, See) und ist bei mehreren Staaten im Einsatz. Pro Jahr werden mindestens 271 Millionen Reisende über dieses System einer Risikoanalyse unterzogen [18]. Die Plattform wurde in praktischen Einsätzen mit unterschiedlichen Regierungen erprobt und kontinuierlich weiterentwickelt, um sich an verändernde Bedrohungslagen anzupassen [18].

Funktionsweise: SITA's PIU-Lösung deckt die gesamte *Reisekette* ab – *vor, an* und *nach* der Grenze [19]. Schon vor der Reise erfolgt eine Vorabprüfung (Pre-Screening) aller Passagiere. Hierbei sammelt die Software API/PNR-Daten sowie weitere verfügbare Quellen (z.B. Visa-Informationen, elektronische Reisegenehmigungen, ggf. offene Quellen) und reichert sie zu umfassenden Personenprofilen an [19]. Durch den Abgleich mit nationalen und internationalen Watchlists, Fahndungs- und No-Fly-Listen identifiziert das System bekannte Personen von Interesse bereits vor der Einreise [18]. Darüber hinaus können die Behörden flexible Risiko-Regeln definieren, um verdächtige Muster zu erkennen – zum Beispiel bestimmte Reiserouten, Zahlungsweisen oder Häufungen von Kurzbuchungen. Diese Threat Profiles lassen sich einfach erstellen und an neue Bedrohungen anpassen [18]. Die Software verwaltet auch sehr umfangreiche Mengen an Zielfprofilen und kann diese laufend optimieren. Während des laufenden Betriebs ermöglicht SITA Intelligence and Targeting eine dynamische Risikobewertung: Neue Informationen (etwa nach einer Kontrolle vor Ort) fließen in den Analysenkreislauf zurück, wodurch die Trefferqualität sukzessive verbessert wird [19]. Das System unterstützt eine Multi-Agency-Zusammenarbeit, d.h. es können verschiedene Sicherheitsbehörden gemeinsam auf die Plattform zugreifen und Erkenntnisse teilen [18]. Bei identifizierten Hochrisiko-Reisenden kann SITA's Lösung in Verbindung mit einem Interactive API-Dienst (wie SITA Advance Passenger Processing) eine Echtzeit-Rückmeldung an die Airline geben, um z.B. das Boarding zu verweigern [20]. Ankommende Passagiere werden durch das System in Risikokategorien eingestuft, sodass Grenzbeamte bei Ankunft gezielt diejenigen herausziehen können, die einer intensiveren Überprüfung bedürfen.

KI-gestützte Analyse: Die SITA-Lösung verwendet umfangreiche Data Analytics, Machine Learning und künstliche Intelligenz, um aus den Rohdaten Prognosen und Entscheidungen abzuleiten [18]. Sie verfügt über advanced risk assessment methods, die den Zeitaufwand für strategische Risikoanalysen erheblich reduzieren [19]. Konkret bedeutet dies: KI-Algorithmen durchforsten die Datenströme und erkennen auffällige Muster oder Anomalien, die auf unbekannte Risiken hindeuten könnten [19]. Gleichzeitig helfen Qualitäts-Algorithmen, die Zahl der False Positives (fälschlich verdächtige Fälle) zu senken, indem sie Profile präziser schärfen [18]. Die Plattform setzt auf vorausschauende Analytik (predictive profiling), um Risiken möglichst früh „an der Wurzel“ zu packen – idealerweise bevor der Passagier überhaupt ein Flugzeug betritt [19]. In der Praxis wird eine Fülle von heterogenen Daten kombiniert: offene Quellen, Trends aus früheren Vorfällen, Reisehistorien und Erkenntnisse aus aktuellen Ermittlungen fließen in die KI-Modelle ein [19]. So entsteht ein iterativer Analysezyklus: neue Operationsergebnisse (Treffer, Festnahmen, etc.) werden ins System zurückgemeldet, was die KI nutzt, um die Profile weiter zu verbessern [19]. SITA betont, dass durch diese Intelligence-Methodik Risiken nicht nur vor der Grenze erkannt werden, sondern auch laufend während der Reise und nachgelagert überwacht werden können [19]. Die KI-Funktionen arbeiten Hand in Hand mit den regelbasierten Komponenten: Die Behörden behalten die Hoheit, indem sie eigene Regeln definieren und die KI zur Musterentdeckung einsetzen. In Summe unterstützt SITA Intelligence and Targeting die Sicherheitskräfte dabei, gezielt gegen bekannte und unbekannte Bedrohungen vorzugehen, die nationale Sicherheit zu stärken und zugleich den legitimen Reiseverkehr zu fördern [18].

Benutzerfreundlichkeit: Als weltweit eingesetzte Lösung legt SITA Wert darauf, dass die Software in unterschiedliche behördliche Umgebungen passt und einfach bedienbar ist. Die Benutzeroberfläche ist mehrsprachig verfügbar und ermöglicht es, komplexe Analyseprozesse visuell zugänglich zu machen [21]. Über ein konfigurierbares Regel-Management können Nutzer ohne Programmieraufwand neue Risiko-Profile oder Alarmkriterien anlegen und bestehende justieren [18]. Ein integriertes *Business Process Management* sorgt dafür, dass Workflows (etwa die Bearbeitung eines Treffers durch verschiedene Stellen) abgebildet werden können [21]. Die Plattform ist zudem offen für Integration: Durch definierte Schnittstellen (APIs) lässt sie sich in die vorhandene IT-Landschaft (Grenzkontrollsysteme, Datenbanken anderer Behörden) einbinden [21]. Datensicherheit und Datenschutz genießen hohe Priorität – die Einhaltung der Datenschutzgesetze (z.B.

DSGVO) ist gewährleistet, Zugriffe sind rollenbasiert steuerbar und alle Aktionen auditierbar [21]. SITA Intelligence and Targeting ist hochskalierbar (>100 Mio. Passagiere) und kann je nach Kundenwunsch on-premise oder als Managed Service betrieben werden [21]. Die Oberfläche stellt umfangreiche Filter- und Recherchefunktionen bereit, damit Analysten schnell bestimmte Personen, Ereignisse oder Reiserouten finden und detailliert untersuchen können [18]. Insgesamt zeichnet sich die Benutzerfreundlichkeit dadurch aus, dass komplexe Daten in einer einzigen Ansicht konsolidiert werden: Die Regierung erhält *"one view of their inbound travelers"*, also eine Gesamtübersicht aller ankommenden Reisenden mit allen relevanten Details [20]. Aufgrund dieser Übersichtlichkeit können Entscheidungen fundiert und zügig getroffen werden, was sowohl die Sicherheit erhöht als auch den Reiseprozess für unauffällige Passagiere beschleunigt [19].

Grafische Darstellungen und Interfaces: SITA's Plattform bietet interaktive Dashboards und Visualisierungstools, um Lagebilder in Echtzeit zu erfassen. So werden beispielsweise Trends pro Flugroute, Airline oder Zeitperiode grafisch aufbereitet (Trendanalysen) [22]. Verdachtsfälle können auf Karten angezeigt werden – etwa die Reiseroute eines Passagiers oder geografische *Hotspots* von Risiken. Die Operationelle Lage an allen Grenzstationen lässt sich in einer zentralen Konsole überwachen. Ein Nutzer kann Passagierlisten (Manifeste) jeder ankommenden Maschine einsehen, Risikobewertungen pro Passagier abrufen und sogar Sitzpläne visualisieren, um z.B. festzustellen, wo potenziell gefährliche Reisende im Flugzeug saßen. Weiterhin ermöglicht die Software die Darstellung von Netzwerkdiagrammen, die Zusammenhänge zwischen verschiedenen Entitäten verdeutlichen – beispielsweise zeigt ein Graph die Verbindungen zwischen mehreren Reisenden über gemeinsame Kontaktpunkte (Adresse, Telefonnummer etc.). Durch solche Visualisierungen können Analytiker leichter komplexe Fälle verstehen und Ergebnisse auch gegenüber Entscheidungsträgern überzeugend präsentieren. SITA liefert zudem Berichts- und Statistikmodule, die beispielsweise die Wirksamkeit von Profiling-Regeln (Trefferquoten, false positives) in Form von Charts darstellen. Insgesamt unterstützt die Oberfläche sowohl die taktische Echtzeit-Überwachung als auch strategische Analysen durch intuitive Visualisierungen.

2.2.4 Weitere Anbieter

Neben den oben fokussierten Lösungen gibt es weitere Anbieter und Systeme, die PIUs unterstützen, wenngleich sie teils einen anderen Umfang haben:

- **WCO Global Travel Assessment System (GTAS):** Eine Open-Source-Software der Weltzollorganisation (WCO) und US Customs and Border Protection. GTAS ermöglicht die automatisierte Risikoanalyse von API/PNR-Daten und wurde entwickelt, um Mitgliedstaaten eine kostenlose Profiling-Lösung bereitzustellen [13]. Das System kann durch Staaten selbst hinter der Firewall betrieben werden und bietet ähnliche Kernfunktionen (Regel-Engine, Watchlist-Abgleich, Trefferanzeige) wie kommerzielle Produkte. Es gilt als Referenzimplementierung, um ICAO-Standards zu erfüllen, und der Quellcode ist offen einsehbar (Transparenz für Datenprozesse) [13].
- **Palantir Gotham:** Eine Datenintegrations- und Analyseplattform, die von einigen Ländern auch für Passagierdaten genutzt wird. Palantir ist zwar kein spezialisiertes PNR-System, kann aber große Datenmengen unterschiedlicher Art in einem Knowledge-Graph zusammenführen. Gotham transformiert heterogene Daten in einen einheitlichen Datenbestand und reichert diesen um erkennbare Objekte (Personen, Orte, Ereignisse) sowie deren Beziehungen an [14]. Durch mächtige Visualisierungstools (Netzwerkgrafiken, zeitliche Abläufe, Karten) und Kollaborationsfunktionen eignet sich Palantir zur tiefgehenden Analyse komplexer Fälle. Mehrere Sicherheitsbehörden können gemeinsam Analysen durchführen und schrittweise ein Intelligence-Bild aufbauen. Palantir wird jedoch als Closed-Source-Produkt lizenziert und erfordert beträchtliche Investitionen, weshalb es primär von Ländern mit hoher technologischer Ausstattung eingesetzt wird.
- **IBM i2 Analyst's Notebook / Intelligence Analysis Platform:** Dies ist eine etablierte Analysesoftware, die vor allem für Ermittlungen und Geheimdienste entwickelt wurde. IBM i2 ermöglicht die graphische Link-Analyse – das Verbinden von Personen, Ereignissen, Orten durch visuelle Netzwerke – und die Auswertung großer strukturierter und unstrukturierter Datenbestände. In Bezug auf PNR kann i2 genutzt werden, um z.B. Verbindungen zwischen Reisenden oder Buchungsmustern darzustellen. Behörden können PNR-Daten aus ihren Systemen in i2 importieren und dort mit anderen Datenquellen (Ermittlungsakten, Kommunikationsdaten etc.) verknüpfen. Die Stärke von i2 liegt in der Visualisierung komplexer Zusammenhänge und in der Unterstützung des Analyseprozesses (z.B. zeitliche Abläufe, soziale Netzwerke). Einige PIUs nutzen i2 als Ergänzung zu ihren Hauptplattformen, um besonders anspruchsvolle Analysen durchzuführen. Als kommerzielles Tool ist auch hier Schulung und Budget nötig, aber die Software gilt als Gold-Standard für kriminalistische Datenanalyse.
- **Travizory:** Ein relativ neuer Anbieter, der eine integrierte API-PNR Targeting System Lösung anbietet, vor allem für kleinere Staaten und neue Digitalkonzepte. Travizory kombiniert PNR- und API-Daten

mit digitalen Reisegenehmigungen (eVisa) und Biometrie. Die cloud-basierte Plattform liefert Echtzeit-Daten der Carrier, KI-gestütztes Profiling und automatische Warnmeldungen [23]. Besonderes Augenmerk liegt auf der biometrischen Verifikation – bei Integration von Travizorys eVisa-System können Fotos/Fingerabdrücke zur Risikobewertung herangezogen werden [23]. Travizory betont *Datenhoheit* (die Daten verbleiben im Land) und Compliance mit ICAO-Standards [23]. Die Benutzeroberfläche ist sehr modern: Sie zeigt Passagierlisten mit Risiko-Scores, Routen auf interaktiven Karten, Sitzpläne und Verbindungen zwischen Reisenden auf einen Blick [23]. Damit wird ein hoher Bedienkomfort erreicht. Travizory's Lösung wird als Software-as-a-Service angeboten und ist bereits in einigen Ländern (v.a. in Afrika) im Einsatz, um die Einreisegenehmigung und PNR-Analyse zu vereinen.

Weitere Anbieter im Umfeld von PIU-Lösungen sind z.B. Collins Aerospace (ARINC), das APIs-Gateways und Datenmanagement für Grenzbehörden liefert, oder Spezialfirmen für Datenfusion und Risikoanalyse. Viele Staaten entwickeln auch eigene Systeme in Zusammenarbeit mit lokalen IT-Dienstleistern, insbesondere um nationale Anforderungen und Datenschutzbestimmungen maßgeschneidert umzusetzen. Diese sind oft weniger allgemein bekannt, leisten aber einen vergleichbaren Dienst im Rahmen der PIU-Aufgaben.

2.2.5 Vergleich der wichtigsten Features

Zum Abschluss werden die drei Hauptlösungen – UNOCT goTravel, WCC HERMES und SITA Intelligence & Targeting – tabellarisch gegenübergestellt:

Merkmal	UNOCT goTravel	WCC HERMES	SITA Intelligence & Targeting
Grundtyp / Bereitstellung	UN-geförderte Software, kostenfrei für Mitgliedstaaten (Installation in On-Premise Rechenzentrum).	Kommerzielle Komplettlösung (On-Premise bei Regierung); schneller Roll-out (1 Tag)	Kommerzielle Enterprise-Lösung (flexibel On-Premise oder Managed Service); seit Jahren global im Einsatz.
Datenintegration	Single Window für API/PNR (EDIFACT, XML, PaxLst usw.); Standard-Schnittstellen zu Airlines via IATA/ICAO; Datenaustauschmodul zw. PIUs optional.	Single Window für API, iAPI und PNR aller Verkehrsarten; Direktanbindung an Airlines (versch. Formate); effizientes Parsing & Konflikterkennung	Verbunden mit >600 Airlines; unterstützt alle IATA/ICAO-Standards (PNRGOV etc.); Multi-Modal (Luft, Land, See); inkl. Echtzeit-Interactive API für Pre-Boarding Checks.
Risikobewertung & Regeln	Regelbasierte Indikatoren (frei konfigurierbar durch Behörden). Watchlist-Abgleich inkl. Interpol-Infos. Schwerpunkt auf bekannte Muster (weitere durch Community-Updates). Manuelle Abfragefunktion für Ermittler.	Regelbasiertes Targeting (flexible Risikoprofile durch PIU definierbar); Abgleich gegen nationale / internationale Watchlists mit fuzzy Namensmatching. Schnelle Suche auch in historischen PNR-Daten (5+ Jahre).	Umfangreiche Risk-Engine: vordefinierte und eigene Regeln; reduziert False Positives durch Qualitätsalgorithmen. Multi-Behörden-Workflows und Feedback-Schleifen zur fortlaufenden Optimierung der Profile.
KI-gestützte Funktionen	(Teilweise) – primär regelbasiert. Network Analysis Modul identifiziert Beziehungen und "unknown unknowns"; kontinuierliche Weiterentwicklung angestrebt (Community-Driven). Derzeit keine voll integrierte ML-Komponente, aber Plattform für zukünftige KI-Funktionen vorbereitet.	Ja: Anomalieerkennung erkennt unbekannte Auffälligkeiten in Reisedaten; KI für Trendanalyse (Prognose von Trefferzahlen für Ressourcenplanung); ML schlägt neue Risk-Profile vor (ergänzend zu Expertenregeln). KI entscheidet nicht autonom, sondern unterstützt Analysten.	Ja: Umfangreiche Nutzung von AI/ML: Predictive Analytics zur Früherkennung von Bedrohungen; Dynamic risk scoring in Echtzeit; intelligente Mustererkennung über vielfältige Datenquellen (inkl. offene Quellen). KI-Modelle mit Feedback-Schleifen, um Trefferqualität stetig zu verbessern.

Benutzeroberfläche & Usability	Zweckmäßige Web-Oberfläche, fokussiert auf Kernfunktionen (geringer Overhead). Einfache Regel- und Watchlistpflege über Formulare. Schulung durch UNOCT inklusive; Community für Erfahrungsaustausch. Anpassbar an nationale Bedürfnisse (z.B. Sprache, Gesetzeskonformität).	Intuitive Web-Oberfläche, von Grenzbeamten mitentwickelt. Geringer Schulungsaufwand: Trefferverfolgung, Kommentare, Regelanlage alles per GUI möglich. One-stop Fallübersichten; 24/7 Support durch WCC. Mehrsprachig und leicht konfigurierbar.	Moderne, modulare UI mit mehrsprachiger Oberfläche. Voll konfigurierbar (Regel-Editor, Workflows). Multi-Behörden-Zugriff, rollenbasiert. Starke Integrationsfähigkeit (APIs) in bestehende Systeme. Datenschutz und Audit-Funktionalität integriert.
Visualisierung & Analyse-Tools	Solide Basis-Visualisierungen: Treffer in Listen und Karten (Vorauswahl vor Eintreffen). Grafische Netzwerk-Analyse zur Beziehungserkennung (Bsp. gleicher Kontakt). Weniger detaillierte Dashboards, aber Export von Intelligence-Berichten möglich.	Heatmaps und Dashboards zur Mustererkennung in Reisedaten. Übersichtliche Trefferlisten pro Flug/Schiff. Optionale grafische Netzwerk-Darstellungen via Schnittstellen zu anderen graphischen Tools. Schwerpunkt auf schneller Liste / Tabellen-Ansicht mit Filter.	Umfangreiche Analytics UI: Echtzeit-Lagebild aller ankommenden Reisen (Fluglisten mit Risk Scores). Graphische Link-Analysen, um Verbindungen zw. Personen aufzudecken. Kartenansichten der Reiserouten, Sitzpläne, Diagramme zur Trendanalyse – alles in einer Plattform.
Besondere Stärken	Kostengünstig & UN-verifiziert – ideal für Staaten mit begrenzten Ressourcen, dabei konform mit internationalen Standards (ICAO, UNSC). Fördert internationale Zusammenarbeit (User Community, Datenmodule) und kontinuierliches Lernen aus Best Practices. Datenhoheit verbleibt vollständig beim Land (UNOCT ohne Datenzugriff).	Schnelle Implementierung und geringe Betriebshürden; preisgekröntes Namensabgleich-Algorithmus (ELISE) für multi-kulturelle Daten; flexibles Regelwerk + erste KI-Integration im Markt. Optimierte für minimalen Einfluss auf Passagierfluss (HERMES = Hermes, Götterbote, sinnbildlich für Geschwindigkeit).	Umfassende, erprobte Lösung mit globaler Reichweite; deckt alle Transportmodi ab; starke KI- und Data-Mining-Komponente; reduziert False Positives, erhöht Effizienz und Kollaboration. Unterstützt Wachstum (Tourismus) und Sicherheit gleichermaßen (Ausbalancierung).

2.3 KI Methode Proof-of-Concept

Im Antrag wurden zwei verschiedene Ansätze für den Einsatz von KI-Methoden im Bereich der PIU erwähnt. Zum einen wurde ein Bayessches Netz in Betracht gezogen, dass von Experten über die Abhängigkeit von Variablen definiert wird, die Wahrscheinlichkeiten werden vom Modell dann gelernt. Ein weiteres Modell besteht aus einem hybriden Ansatz, in dem auch die Netzwerkstruktur gelernt wird.

Beide Ansätze wurden ohne das spezielle Hintergrundwissen der Problematik der Arbeitsweise der PIU gewählt. Nach Beginn des Projekts haben wir in vielen Meetings die genaue Arbeitsweise der PIU kennengelernt. Von den in Kapitel 3.1 beschriebenen unterschiedlichen Arbeitsaufgaben sind die ersten zwei relativ einfachen Identifizierungen über Namen und Geburtsdatum gegeben. Auch wenn es hier kleinere Optimierungspotenziale gibt, arbeiten die Methoden recht effizient.

Anderes sieht es allerdings bei der Kriterienabfrage aus, da hier die Erstellung recht aufwändig ist, besteht hier das größte Potenzial zur Steigerung der Effizienz. Es wurde aber auch schnell klar, dass es schwierig ist, feststehendes Expertenwissen zum Lernen für diesen Abfragetyp zu bekommen. Typischerweise werden die Methoden in bestimmten Bereichen der Kriminalität (Drogenschmuggel, Geldwäsche durch Geldtransport, Menschenhandel etc.) sehr schnell an die Zollkontrollen angepasst. Bemerkt z.B. ein Zollbeamter, dass ein

Typus von bestimmten Personen (und entsprechenden Begleitumständen) im Gepäck Geld schmuggelt, wird dieses Wissen in eine Kriterienabfrage umgesetzt. Diese benötigt aber eine feine Optimierung. Werden „zu viele“ Treffer generiert, enthalten diese eine hohe Zahl von Falschverdächtigungen, was kontraproduktiv ist. Zudem muss diese deutlich erhöhten Treffen dann auch wirklich kontrolliert werden, was aber schwer durch das vorhandene Personal umgesetzt werden kann. D.h. es ist eine sehr feingliedrige Analyse notwendig, um zielgerichtete Abfragen zu erstellen, die wenige Haupttreffern mit sehr wenigen Falschtreffern erreichen. In Folge werden einige Personen kontrolliert und abgefangen. Nach einige Wochen merken das naturgemäß auch die dahinterstehenden kriminellen Organisationen. Diese verändern dann ihre Methoden, um wieder „erfolgreich“ tätig zu sein.

Deswegen scheint es vielversprechender zu sein, die Erstellung von zielgerichtete Kriterienabfragen zu vereinfachen und damit zu beschleunigen.

So wurde die Idee entwickelt, dass mit KI-Methoden die Erstellung dieser Abfragen deutlich vereinfacht wird. Dabei kommen Large Language Modelle (LLMs) zum Einsatz, die eine Erstellung der Abfrage mittels natürlicher Sprache ermöglichen. Dieses soll in Form eines Chat-Bots ermöglicht werden, in dem die Abfragen Schritt für Schritt verfeinert wird und dabei die Abfrageergebnisse bei jeder Verfeinerung aktualisiert werden. Zudem kann es ermöglicht werden, dass bei zu hohen Trefferquoten leicht ein Schritt zurückgegangen werden, um die Auswirkungen zu analysieren. Bei falscher Auswahl der Kriterien muss nicht wieder von vorne anfangen werden. In dem man die Möglichkeit vorsieht, diese Abfrage auf eine Testdatenbank anzuwenden, kann schon bei der Erstellung der Abfrage die Trefferzahl und damit die Einsatzfähigkeit abgeschätzt und optimiert werden.

Diese Idee wurde dann prototypisch in AP 3 demonstriert und vom BMI als äußerst vorteilhaft angesehen. Im Kapitel 5 ist die Prototypische Umsetzung dargestellt.

3 ANONYMES BIOMETRISCHES MATCHING

3.1 Einleitung

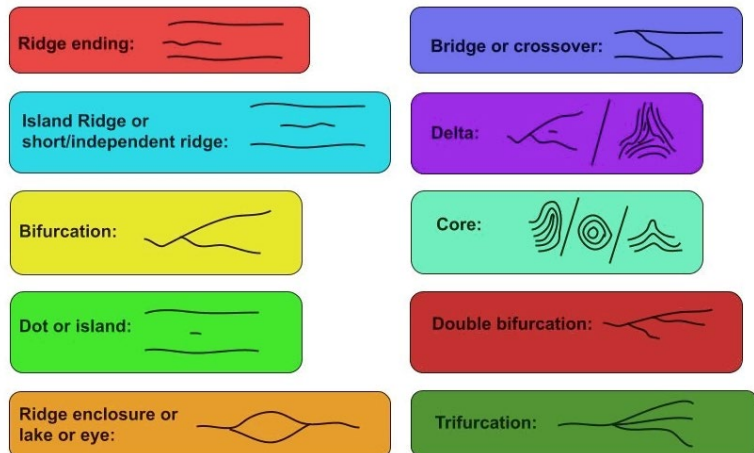
Biometrische Verfahren wie die Fingerabdruckerkennung sind heute weit verbreitet zur Identifikation und Authentifizierung von Personen. Fingerabdrücke sind einzigartig und lebenslang unveränderlich, was sie grundsätzlich zu zuverlässigen Merkmalen macht. Allerdings führt gerade diese Permanenz zu erheblichen Datenschutzbedenken: Geht ein Fingerabdruck einmal verloren oder wird kompromittiert, kann er – anders als ein Passwort – nicht einfach widerrufen oder geändert werden [24]. Unkontrollierte Speicherung und Verarbeitung von Fingerabdrücken gefährden daher die Privatsphäre, da Missbrauch oder Datenlecks die Identität des Betroffenen dauerhaft kompromittieren könnten. Ziel des *anonymen biometrischen Matchings* ist es deshalb, Fingerabdrücke vergleichen zu können, ohne die Roh-Biometriedaten oder die Identität der beteiligten Personen preiszugeben.

In der Praxis bedeutet dies, dass z.B. bei einer Zutrittskontrolle oder einem Datenbankabgleich überprüft werden kann, ob ein vorgelegter Fingerabdruck mit einem gespeicherten Fingerabdruck übereinstimmt, ohne dass der Dienstanbieter jemals den Fingerabdruck im Klartext sieht. Verschiedene kryptographische und systemarchitektonische Ansätze werden hierfür erforscht. Zu den wichtigsten zählen Homomorphe Verschlüsselung, Secure Multi-Party Computation (MPC) – zu Deutsch sichere Mehrparteienberechnung – sowie Distributed-Ledger-Technologien (Blockchains). Diese Methoden versprechen, hochsensible biometrische Merkmale nur in verschlüsselter oder verteilter Form zu verarbeiten, sodass kein einzelner Akteur die vollständigen Klartextdaten erhält.

Der folgende Bericht gibt zunächst einen Überblick über diese Methoden und erläutert deren Funktionsprinzip im Kontext der Fingerabdruckerkennung. Anschließend werden konkrete Umsetzungen aus aktuellen Forschungsarbeiten vorgestellt – insbesondere drei Publikationen von *Julia Mader et al.*, die sich im Rahmen dieses Projekts intensiv mit Privacy-Preserving Fingerprint Matching befassen. Darauf aufbauend werden die Vor- und Nachteile der Methoden diskutiert und bewertet, inwieweit sie sich für anonymes biometrisches Matching eignen. Abbildung 1 zeigt zur Einstimmung einige typische Minutien (Merkmale) in Fingerabdruckmustern, die bei der Matcherstellung eine Rolle spielen.

Typische Minutienmuster in Fingerabdrücken (schematische Darstellung). Fingerabdrücke weisen charakteristische Rückenlinien-Endungen und Verzweigungen (Bifurkationen) auf, die als Minutien extrahiert werden (Darstellung aus [25], CC BY-SA). Diese Minutien werden in sogenannten Templates gespeichert und bilden die Basis für den Vergleich der Fingerprints einer Person die mit verifizierten Originalen verglichen werden um die Identität zu verifizieren. Derzeit werden bei Reisepässen allerdings Fotos von Fingerprints in einer verschlüsselten Form auf dem Chip im Pass gespeichert, und können bei Einreise in ein anderes Land (aufgrund möglicher bilateraler Abkommen) ausgelesen und verglichen werden. Das bedeutet aber auch, dass zwischen den Einreiseland und dem Aussteller Land des Passes ein gewisses Vertrauen existieren muss, diese Daten nicht zweckwidrig zu verwenden.

MINUTIAE PATTERNS



3.2 Methoden für anonymes biometrisches Matching

3.2.1 Homomorphe Verschlüsselung (HE)

Homomorphe Verschlüsselung erlaubt Berechnungen direkt auf verschlüsselten Daten, ohne dass diese zuvor entschlüsselt werden müssen. Im Kontext der Fingerabdruckerkennung bedeutet das: Ein Fingerabdruck wird auf Seiten des Besitzers verschlüsselt und an einen Server übertragen; der Server kann dann einen Abgleich mit gespeicherten, ebenfalls verschlüsselten Fingerabdruck-Templates durchführen, ohne jemals Zugriff auf die Klartext-Merkmale zu haben [26]. Das Ergebnis der Berechnung – z.B. eine Ähnlichkeitskennzahl – liegt wiederum verschlüsselt vor und wird erst beim Berechtigten entschlüsselt. Auf diese Weise bleiben die biometrischen Daten während des gesamten Prozesses vertraulich. Dieses Verfahren bietet den Vorteil, dass kein vertrauenswürdiger Dritter benötigt wird; ein einzelner Server kann den Abgleich ausführen, da die Sicherheit allein auf der mathematischen Stärke der Verschlüsselung beruht. Selbst ein kompromittierter Server käme ohne den geheimen Schlüssel nicht an die Fingerabdruckdaten selber heran [26].

In der Praxis stehen für homomorphe Berechnungen zwei Ansätze zur Verfügung: *Teilhomomorphe* Verschlüsselungsverfahren (die nur bestimmte Operationen wie Addition oder Multiplikation auf Ciphertexts erlauben) oder *vollhomomorphe* Verfahren (die im Prinzip beliebige Berechnungen unterstützen). Da ein Fingerabdruckvergleich komplexe Algorithmen umfasst (z.B. das Auffinden und Abgleichen von Dutzenden Minutienpaaren), ist meist ein vollhomomorphes Verschlüsselungsschema erforderlich, um einen echten minutienbasierten Matcher umzusetzen. Allerdings ist die Vollhomomorphe Verschlüsselung auch sehr rechenaufwändig. Ein aktuelles Beispiel zeigt die Größenordnung: Kim *et al.* (2020) implementierten ein Fingerabdruck-1:1-Matching mit dem TFHE-Framework (vollhomomorphe Verschlüsselung über Torus) – die Berechnung eines einzelnen Vergleichs dauerte im Schnitt rund **166 Sekunden** [27]. In sicherheitskritischen Anwendungen (z.B. Zugangskontrolle in Echtzeit) ist eine solche Verzögerung kaum akzeptabel.

Um die Berechnungen zu vereinfachen, setzen einige HE-Ansätze auf vereinfachte Fingerabdruck-Templates. Anstelle aller Minutiendetails wird etwa ein fester Feature-Vektor (z.B. ein sogenannter *Fingercode*) aus dem Fingerabdruck abgeleitet, der dann homomorph verglichen wird (etwa mittels euklidischer Distanz) [24], @Yang2023. Das reduziert die Komplexität und konnte bereits früh praktikable Laufzeiten erzielen [24]. Allerdings leidet darunter oft die Genauigkeit des Matchings. Ein Vergleich von Verfahren zeigt, dass ein featurevektor-basierter Ansatz (hier: Fingercode-Matcher) Fehlerraten (EER) von 18–30 % aufweist, während ein minutienbasierter Algorithmus (SourceAFIS) auf EER $\approx 7,9\%$ kommt [24]. Mit anderen Worten: Einfache Vektordarstellungen führen zu signifikant mehr Fehlentscheidungen – in der Praxis ein untragbarer Nachteil. Moderne Forschung zielt daher darauf ab, trotz homomorpher Verarbeitung die *Detailtiefe* eines minutienbasierten Matchers zu erhalten, was jedoch anspruchsvoll ist.

Vorteile: Homomorphe Verschlüsselung bietet ein sehr hohes Datenschutzniveau, da Daten zu jedem Zeitpunkt verschlüsselt bleiben. Ein einzelner Server oder Cloud-Dienst kann den Abgleich vornehmen, ohne dass dem Betreiber vertraut werden muss – ein attraktives Szenario für Cloud-basierte biometrische

Identitätsdienste. Zudem ist keine Aufteilung der Geheimnisse auf mehrere Stellen erforderlich, was die Systemarchitektur vereinfacht.

Nachteile: Der große Nachteil sind die hohen Rechen- und Zeitkosten. Selbst mit Optimierungen sind homomorphe Fingerabdruck-Matcher derzeit (Stand 2024) nicht in Echtzeit einsetzbar, wenn die volle Genauigkeit eines Minutenabgleichs erreicht werden soll. Zudem müssen ggf. Anpassungen an den Matching-Algorithmen vorgenommen werden (z.B. Nutzung eines festlängen-Codes statt variabler Minutenlisten), was die Erkennungsleistung verschlechtern kann [24]. Forschungsergebnisse zeigen zwar die Machbarkeit (Sicherheit) solcher Verfahren, aber die Effizienz bleibt eine Herausforderung [27]. Fortschritte in der Kryptographie und Hardware (z.B. Spezialchips für homomorphe Operationen) könnten diese Einschränkungen in Zukunft reduzieren.

3.2.2 Secure Multi-Party Computation (MPC)

Ein alternativer Ansatz ist die Sichere Mehrparteienberechnung (Secure MPC). Hierbei wird die Berechnung auf mehrere Parteien bzw. Server verteilt, so dass keine einzelne Instanz sämtliche Daten im Klartext erhält. Das klassische Szenario: Der biometrische Input (z.B. die extrahierten Minuten eines Fingerabdrucks) wird in sogenannte *Geheimanteile* (Shares) zerlegt und auf mehrere unabhängigen Server verteilt. Ebenso werden die gespeicherten Fingerabdruck-Templates aufgeteilt. Die Server führen dann einen kryptographischen Protokollablauf aus, bei dem sie interaktiv eine Funktion berechnen, ohne ihre jeweils geheimen Anteile preiszugeben [24]. Am Ende kennen die Server nur das Berechnungsergebnis (z.B. „Match“ oder „No Match“), nicht aber die zugrundeliegenden Fingerabdrücke. Solche MPC-Protokolle basieren z.B. auf Secret Sharing und bestimmten Operationen (Addition, Multiplikation) im geteilten Geheimnis oder auf Yao-Garbled-Circuits – je nach Ausgestaltung.

Im Unterschied zur homomorphen Verschlüsselung, bei der ein einzelner Server auf Ciphertexts rechnet, erfordert MPC also eine verteilte Architektur mit N Servern (typischerweise 2 oder 3 in praktikablen Systemen). Der Sicherheitsgewinn entsteht durch die Annahme, dass nicht alle Server kollaborieren, d.h. solange mindestens ein Server ehrlich ist und nicht mit einem Angreifer zusammenspielt, bleiben die Daten geheim. MPC eignet sich gut, um komplexe Algorithmen abzubilden, da prinzipiell jede Berechnung in logische Grundoperationen zerlegt und sicher ausgeführt werden kann. Allerdings steigt mit der Komplexität auch hier der Rechen- und Kommunikationsaufwand beträchtlich. MPC hatte lange den Ruf, „zu langsam“ für praktische Anwendungen zu sein [24]. Neuere Entwicklungen – optimierte Protokolle, spezialisierte Frameworks und Hardware – haben diese Latenz jedoch drastisch senken können [24].

Gerade im Bereich Fingerabdruckmatching gibt es eindrucksvolle Fortschritte. Mader und Lorünser (2024) demonstrierten einen voll minutenbasierten SourceAFIS-Matcher unter MPC-Bedingungen. Durch Optimierungen und die Nutzung des effizienten MPC-Frameworks MP-SPDZ erzielten sie eine Vergleichszeit von unter 7 Sekunden pro 1:1 Abgleich, während frühe Prototypen noch *Stunden* benötigten [24]. Damit rückt eine Echtzeit-Fähigkeit in greifbare Nähe. Gleichzeitig blieb die Genauigkeit auf dem Niveau des Originalalgorithmus (FNMR ca. 2,5 % bei FMR 0,1 % laut SourceAFIS-Auswertung) [24]. Diese Resultate „zeigen das *praktische* Potenzial moderner MPC-Techniken für privatsphäreschonendes Fingerabdruck-Matching“ [24].

Ein großer Vorteil von MPC ist, dass sich Mehrparteien-Szenarien realisieren lassen, in denen verschiedene Institutionen jeweils Teilinformationen beitragen, ohne ihre Daten vollständig offenlegen zu müssen. Ein oft genanntes Beispiel ist der Abgleich eines Reisenden an der Grenze gegen internationale Fahndungslisten (*No-Fly-Listen*). Traditionell müssten alle Beteiligten ihre Fingerabdruckdaten zentral zusammenführen – was aus Datenschutzgründen problematisch ist. Mit MPC kann stattdessen jedes Land seine Liste in verschlüsselter Form beisteuern (über mehrere Server verteilt); der Fingerabdruck des Reisenden wird ebenfalls in Anteile aufgeteilt und an die Server gesendet. Gemeinsam prüfen die Server nun, ob ein Match vorliegt, und geben nur im Trefferfall eine Meldung aus [24]. Die Privatsphäre unauffälliger Reisender bleibt gewahrt (ihre biometrischen Daten werden nie entschlüsselt weitergegeben), und gleichzeitig wird durch die verteilte Datenhaltung verhindert, dass eine zentrale Stelle alle No-Fly-Daten einsehen kann [24]. Dieses Szenario zeigt die Stärke von MPC bei gleichzeitiger Wahrung von Sicherheit *und* Privatsphäre in kollaborativen Anwendungen.

Vorteile: MPC erlaubt die Verarbeitung komplexer Matching-Algorithmen (z.B. Minutenvergleich mit vielen Teilschritten) unter Wahrung der Vertraulichkeit. Anders als bei HE muss der Algorithmus nicht drastisch vereinfacht werden – die volle Erkennungsleistung bleibt erhalten. Zudem können verschiedene Datenhalter zusammenarbeiten, ohne ihre sensiblen Daten offenlegen zu müssen. Durch verteilte Verantwortung ergibt sich potentiell auch höhere Ausfallsicherheit. Dank intensiver Forschung (optimierte Protokolle wie SPDZ, ABY, etc.) sind die Performanceeinbußen in vielen Fällen auf ein praktikables Maß reduziert worden [24]. Bei Mader *et al.* wurde etwa durch spezifische Optimierungen (Entschärfung von MPC-unfreundlichen Schleifen,

Parallelisierung, selektives Klarhalten nicht-sensibler Indexwerte) die Laufzeit um den Faktor 1500 verkürzt ([24]).

Nachteile: Der Hauptnachteil ist der Systemaufwand: Es werden mehrere unabhängige Server oder Instanzen benötigt, die idealerweise in unterschiedlichen Verwaltungsdomänen liegen (um Kollusion zu verhindern). Die Koordination dieser Parteien und die Netzwerkkommunikation machen das System komplexer als einen Single-Server-Ansatz. Zudem setzt die Sicherheit voraus, dass *mindestens* eine Partei ehrlich agiert – bei absichtlicher Absprache aller Beteiligten könnten auch in MPC vertrauliche Daten offengelegt werden. In der Praxis ist dieses Szenario durch organisatorische Maßnahmen abzumildern (z.B. Betrieb der Server durch unterschiedliche Behörden oder Unternehmen). Performance-technisch erzeugt MPC immer noch einen Overhead gegenüber ungeschütztem Matching, wenngleich dieser – wie gezeigt – mittlerweile stark gesunken ist. Schließlich erfordert die Implementierung eines komplexen Matchers in einem MPC-Framework erhebliche Entwicklungsarbeit und Expertise in Kryptographie, was die Eintrittshürde erhöht.

3.2.3 Distributed Ledger / Blockchain

Neben den rein kryptographischen Verfahren wird auch der Einsatz von Distributed-Ledger-Technologien (wie z.B. Blockchains) für biometrische Anwendungen erforscht. Eine Blockchain ist im Kern ein dezentral geführtes, unveränderliches Register von Datensätzen, das auf vielen Knoten redundant gespeichert wird [28]. In Zusammenhang mit Fingerabdruck-Identitäten verspricht diese Technologie vor allem zwei Dinge: Integrität und Dezentralisierung. Erstens kann ein einmal auf der Blockchain gespeichertes biometrisches Merkmal (bzw. dessen Hash oder verschlüsseltes Template) nicht unbemerkt manipuliert werden – die Historie ist auditierbar und fälschungssicher. Zweitens gibt es keine zentrale Kontrollinstanz; die Datenhoheit liegt verteilt, was Single-Point-of-Failure und Monopole verhindert [28].

Für anonymes Matching ist eine Blockchain jedoch nicht als alleinige Lösung ausreichend, sondern eher als Komponente in einem Gesamtsystem. Typischerweise würde eine Blockchain genutzt, um *Referenzwerte* oder Berechtigungsnachweise abzulegen: Beispielsweise könnte jeder Benutzer einen verschlüsselten Fingerabdruck-Template oder einen biometrischen Hash auf der Blockchain hinterlegen, verknüpft mit einer digitalen Identität (DID – Decentralized ID). Bei einer Authentifizierung könnte der Nutzer dann seinen aktuellen Fingerabdruck vorlegen und in einem Protokoll (ggf. mittels der oben beschriebenen HE- oder MPC-Methoden) nachweisen, dass sein Fingerabdruck mit dem registrierten Wert auf der Blockchain übereinstimmt – ohne diesen selbst offen zu legen. Die Blockchain fungiert hier als vertrauenswürdiger Speicherort für die biometrischen Referenzdaten, der von keinem einzelnen Provider manipuliert werden kann [28]. Insbesondere in Self-Sovereign-Identity-Ansätzen ist dies attraktiv: Benutzer kontrollieren ihre eigenen Daten und legen nur Prüfsummen oder verschlüsselte Tokens offen, die global verifizierbar sind.

Die Verwendung einer Blockchain bringt dabei Stärken und Schwächen mit sich. Auf der Pro-Seite steht die bereits erwähnte Manipulationssicherheit und Transparenz. Alle Transaktionen (z.B. Registrierungen oder erfolgreiche Verifikationen) können lückenlos nachvollzogen werden. Für den Datenschutz ist wichtig, dass biometrische Daten auf der Chain immer nur in geschützter Form (verschlüsselt oder gehasht) gespeichert werden, da eine Blockchain per se öffentlich oder zumindest vielen Parteien zugänglich ist [28]. Werden diese Schutzmaßnahmen eingehalten, kann die dezentrale Speicherung sogar das Missbrauchsrisiko reduzieren – es gibt kein zentrales Datenleck, das ein Angreifer kompromittieren kann [28]. Ein weiterer Vorteil ist die mögliche Automatisierung via Smart Contracts: So ließen sich z.B. Zugriffsfreigaben an biometrische Daten automatisiert steuern. Ein Beispiel: Ein Smart Contract könnte regeln, dass eine verschlüsselte Fingerabdruckvorlage nur dann zur Prüfung freigegeben wird (durch Entschlüsselung via MPC oder ähnliches), wenn zwei unabhängige Berechtigte zugestimmt haben. Blockchain und Smart Contracts können also als *Orchestrierungs- und Vertrauensschicht* dienen.

Vorteile: Distributed Ledgers bieten *Verifizierbarkeit* und *Vertrauenswürdigkeit* ohne zentrale Autorität. Für die Nutzer entsteht potentiell ein höheres Gefühl der Kontrolle und Sicherheit ihrer Daten, da Missbrauch durch einzelne Dienstleister erschwert wird [29]. In Kombination mit biometrischen Verfahren kann Blockchain dazu beitragen, dass Transaktionen (z.B. der Identitätsnachweis) fälschungssicher dokumentiert werden und kein zentraler Biometriespeicher mehr nötig ist – was z.B. Regierungsprojekten die Einhaltung strenger Datenschutzgesetze erleichtern könnte. Zudem können durch die globale Verfügbarkeit einer öffentlichen Blockchain biometrische *Authentifizierungen nahezu in Echtzeit* über Grenzen hinweg überprüft werden (theoretisch kann jeder Knoten eine Prüfung starten), ohne klassische Vermittlerstellen [29].

Nachteile: Auf der Contra-Seite stehen Datenschutz- und Performancefragen. Eine öffentliche Blockchain ist transparent – auch verschlüsselte biometrische Daten könnten Angriffspunkte bieten, etwa wenn kryptographische Verfahren in Zukunft gebrochen werden: Dann lägen hochsensible Daten unwiderruflich offen, da sie ja nicht gelöscht werden können (Immutabilität). Außerdem sind aktuelle Blockchain-Netzwerke nicht dafür ausgelegt, große Datenmengen (wie biometrische Templates) effizient zu speichern oder komplexe Berechnungen darauf auszuführen. In der Regel müsste die eigentliche Matching-Berechnung *Off-Chain*

erfolgen (z.B. mittels HE/MPC, siehe oben) und nur das Ergebnis oder ein Proof on-chain abgelegt werden. Dies erhöht die Systemkomplexität. Weiterhin bringt die Dezentralität Verzögerungen durch Konsensfindung mit sich – ein Echtzeit-Abgleich könnte unter Umständen durch die Latenz des Ledgers limitiert sein. Schließlich ist die junge Integration von Blockchain und Biometrie noch mit offenen Fragen behaftet, etwa wie Widerruf und Anonymität genau gewahrt werden. Zwar kann man Pseudonyme verwenden, doch biometrische Identitäten sind per se eindeutig – hier müssen zusätzliche Mechanismen (z.B. Zero-Knowledge-Beweise) eingesetzt werden, um Nutzeraktivitäten auf der Chain nicht deanonymisierbar zu machen. Diese Themen werden in aktuellen Forschungsarbeiten (u.a. rechtliche Analysen und SWOT-Studien) intensiv diskutiert [29].

Zusammenfassend lässt sich sagen, dass Distributed Ledgers eine ergänzende Technologie für anonymes biometrisches Matching darstellen, die insbesondere die vertrauenswürdige Verwaltung und Verteilung der Daten adressiert. Für das eigentliche Matching werden sie voraussichtlich mit kryptographischen Verfahren kombiniert: die Blockchain als sicheres Register, homomorphe Verschlüsselung oder MPC für die vertrauliche Berechnung. Erste Pilotprojekte und Studien deuten darauf hin, dass eine solche Kombination praktikable und sichere Lösungen ermöglichen kann [29], allerdings steht ein breiter praktischer Einsatz noch aus.

3.3 Umsetzung in aktuellen Forschungsarbeiten (Fallstudien)

Im Folgenden werden kurz die drei Publikationen, die im Rahmen des Projekts erstellt wurden, betrachtet. Diese vorgestellten Methoden werden im Kontext *Privacy-Preserving Fingerprint Matching* untersucht und implementiert. Alle drei Arbeiten stammen von Julia Mader und Kollegen und zeichnen ein Bild des Fortschritts auf diesem Gebiet zwischen 2023 und 2024.

(1) Masterarbeit von Julia Mader (2023): "Privacy-Preserving Fingerprint Matching". In ihrer MSc-Arbeit legte Mader den Grundstein für die nachfolgenden Studien. Die Arbeit analysiert ausführlich die Sicherheitsrisiken konventioneller Fingerabdrucksysteme und diskutiert mögliche Gegenmaßnahmen, darunter homomorphe Verschlüsselung und Mehrparteienberechnung. Ein Schwerpunkt lag auf dem Entwurf eines *privacy-preserving* Protokolls für den Minutienvergleich. Mader untersuchte dabei die *Machbarkeit* verschiedener Ansätze unter den Gesichtspunkten Sicherheit, Genauigkeit und Effizienz. Wahrscheinlich zeigte die Arbeit zunächst die Grenzen rein homomorpher Ansätze auf (vgl. Laufzeitproblem bei vollhomomorpher Berechnung) und wählte dann MPC als vielversprechende Lösung. In Zusammenarbeit mit dem Austrian Institute of Technology wurde ein Prototyp entwickelt, der die Fingerabdruck-Features in verteiltem Zustand verarbeitet. Die Masterarbeit kam zum Schluss, dass ein minutienbasierter Abgleich möglich ist, ohne biometrische Daten im Klartext offenzulegen, jedoch nur mit speziellen Optimierungen und geeigneter Kryptounterstützung. Diese Erkenntnisse flossen direkt in die nächste Publikation ein.

(2) Mader & Lorünser (ICISSP 2024): "Feasibility of Privacy Preserving Minutiae-Based Fingerprint Matching". Diese Konferenzpublikation präsentiert die erste voll funktionsfähige Implementierung eines Minutien-Matchers auf MPC-Basis. Ausgangspunkt war die Open-Source-Bibliothek *SourceAFIS*, ein minutienbasiertes Fingerabdruckvergleichsverfahren. Mader & Lorünser integrierten *SourceAFIS* in das MPC-Framework MP-SPDZ und nahmen spezifische Anpassungen vor, um die Performance zu steigern. So blieben z.B. die Indexierung von Minutien und Kanten teilweise unverschlüsselt, während sensible Informationen wie Koordinaten und Winkel strikt geheim gehalten wurden – ein Kompromiss, der die Rechenlast reduzierte, ohne die Privatsphäre zu gefährden. Die Leistungsbewertung zeigte eindrucksvolle Ergebnisse: Die Autoren berichten, dass ein einzelner 1:1-Abgleich in unter 10 Sekunden abgeschlossen werden konnte – konkret lag die Zeit *unter 7 Sekunden* im Durchschnitt. Dies ist insofern bemerkenswert, als frühere Versuche (z.B. ein Ansatz mit dem MPyC-Framework) noch Stunden pro Vergleich benötigt hatten. Damit wurde erstmals demonstriert, dass Privacy-Preserving Fingerprint Matching in einem für praktische Anwendungen vertretbaren Zeitrahmen möglich ist. Auch die Genauigkeit blieb hoch: Mit dem MPC-gestützten *SourceAFIS* erreichten sie ähnliche Fehlerraten wie das Original (z.B. EER um 7–8 %). Die Arbeit diskutiert zudem die Skalierbarkeit für 1:N-Abgleiche und konstatiert, dass für datenbankweite Suchen (Identifikation) weitere Optimierungen nötig sein würden – doch für Authentifizierungsfälle (1:1-Vergleiche) sei die gezeigte Lösung bereits vielversprechend. Insgesamt leistet diese Publikation einen Machbarkeitsnachweis: Sie untermauert, dass MPC trotz seines Rufes als langsam inzwischen reif ist, einen echten Fingerabdruckvergleich sicher zu realisieren. Als Ausblick regen Mader & Lorünser an, die Ansätze auf größere Datenbanken und strengere Echtzeitanforderungen auszuweiten, was direkt zur dritten Arbeit führt.

(3) Mader et al. (WPES 2024): "Towards Real-Time Privacy-Preserving Minutiae Matching". In dieser aktuellsten Veröffentlichung gehen Mader und Kollegen den nächsten Schritt und nähern sich dem *Echtzeit*-Ziel noch weiter an. Aufbauend auf dem zuvor entwickelten System, konzentriert sich diese Arbeit auf zwei Aspekte: Performance-Optimierung und Verbesserung der Erkennungsgenauigkeit. Zum einen wurden zusätzliche Feinoptimierungen im MP-SPDZ-basierten Protokoll vorgenommen, um die Berechnungen parallelisierbarer und die Kommunikationsrunden effizienter zu machen. Zum anderen integrierte das Team verschiedene Minutien-Extractoren in das System. Das bedeutet, es wurden mehrere Methoden ausprobiert,

um aus dem Fingerabdruck die Merkmalsvektoren zu gewinnen – z.B. alternative Filter oder Qualitätsverbesserungen der Minutiendaten –, und diese Varianten wurden unter MPC-Bedingungen getestet. Durch die Wahl eines ggf. robusteren oder präziseren Extraktors ließen sich die Fehlerkennzahlen deutlich senken. Mader *et al.* berichten von „*signifikanten Verbesserungen der Error Rates*“, was darauf hindeutet, dass die False-Match-Rate und False-Non-Match-Rate gegenüber dem vorherigen Ansatz weiter reduziert wurden. Genaue Zahlen werden im Abstract nicht genannt, aber es ist plausibel, dass die EER Richtung $<5\%$ gedrückt werden konnte. Die Arbeit zeigt damit, dass man nicht allein die Kryptographie betrachten darf, sondern auch die vorgelagerte Signalverarbeitung optimieren kann, um trotz Datenschutz die biometrische Leistungsfähigkeit hochzuhalten. Hinsichtlich Laufzeit nähert man sich dem Echtzeit-Bereich: Der Titel „*Towards Real-Time*“ signalisiert, dass das System kurz davorsteht, unter üblichen Bedingungen (Netzwerk, Server) so schnell zu sein wie ein konventioneller Matcher. Möglicherweise erreichen sie Abgleichzeiten im Bereich weniger Sekunden bis unter 1 Sekunde – genaue Werte sind hier nicht verfügbar, aber das Ziel ist klar die $<1s$ Marke. Insgesamt *unterstreichen die Ergebnisse die Machbarkeit und Effektivität moderner MPC-Techniken für privates Fingerabdruck-Matching*. Die Autoren betonen, dass weitere Optimierungen (z.B. durch noch stärkere Parallelisierung oder spezialisierte Hardware) den Ansatz bald *einsatzreif für echte Echtzeitanwendungen* machen könnten. Diese Arbeit repräsentiert den aktuellen Stand der Technik: ein System, das sowohl sicher (anonym), genau als auch nahezu in Echtzeit ist.

Die drei Publikationen zeigen in ihrer Entwicklung eine klare Tendenz: Weg von theoretischen Betrachtungen hin zu praktischen Implementierungen, welche die Lücke zwischen Sicherheit und Nutzbarkeit schließen. Während die Masterarbeit das Problem und Lösungsideen umfassend beleuchtet, fokussieren die Konferenzbeiträge auf die MPC-Lösung, da diese sich als effizientester Kompromiss herausgestellt hat. Mader *et al.* liefern damit einen Proof-of-Concept, der über das rein Akademische hinausgeht – ein wichtiges Signal dafür, dass anonymes biometrisches Matching kein utopisches Vorhaben mehr ist, sondern mit heutigem Stand der Technik machbar. Besonders hervorzuheben ist, dass alle Arbeiten Open-Source-Tools (SourceAFIS, MP-SPDZ) nutzen und weiterentwickeln, was die Reproduzierbarkeit und Transparenz fördert. Auch wurden reale Fingerabdruckdatensätze (z.B. FVC2002/BioSec) in den Experimenten verwendet, um praxisnahe Fehlerraten und Zeiten zu ermitteln. Kritisch anmerken kann man, dass das gezeigte System bislang auf *1:1 Vergleiche* optimiert ist. In vielen Anwendungen (z.B. Polizeiliche Identifizierung) muss jedoch *1:N Matching* erfolgen, was die Datenmenge und damit die MPC-Kosten massiv erhöht. Hier stehen belastbare Ergebnisse noch aus; die Autoren verweisen aber selbst auf diese Herausforderung und planen, in zukünftigen Arbeiten auch dieses Problem anzugehen. Nichtsdestotrotz sind die präsentierten Fallstudien ein eindrucksvoller Beleg dafür, wie Theorie (Kryptographie) und Praxis (Biometrie) vereint werden können, um Datenschutz und Sicherheit gleichzeitig zu gewährleisten.

3.4 Bewertung der Methoden: Eignung für anonymes biometrisches Matching

Nach der detaillierten Betrachtung der Methoden und Veröffentlichung soll nun vergleichend bewertet werden, wie gut sich Homomorphe Verschlüsselung, MPC und Distributed Ledgers für anonymes biometrisches Matching eignen. Dabei werden Schutzgrad, Leistungsfähigkeit, Komplexität und Einsatzszenarien gegenübergestellt.

- **Datenschutz und Sicherheit:** Alle drei Ansätze erhöhen den Datenschutz im Vergleich zur trivialen Lösung (direkter Klartextvergleich auf einem Server). Homomorphe Verschlüsselung bietet dabei theoretisch den stärksten Schutz, da *weder während der Übertragung noch während der Verarbeitung* jemals Klartextdaten anfallen – die sensiblen Merkmale bleiben mathematisch abgeschirmt [26]. MPC kommt dem sehr nahe; hier fallen die Klartextdaten verteilt an, was bei genügend vielen unabhängigen Parteien einem ähnlich hohen Schutz gleichkommt. Allerdings besteht ein Restrisiko bei MPC, falls alle beteiligten Server kolludieren – dieses Szenario ist unwahrscheinlich, aber nicht unmöglich. Distributed Ledger per se schützt Daten nicht kryptographisch (jeder Blockchain-Knoten könnte Daten lesen), sondern setzt auf *Transparenz und Unveränderbarkeit*. In Kombination mit Verschlüsselung kann aber auch hier ein sehr hohes Schutzniveau erreicht werden [28]. Betrachtet man gezielte Angriffe: Ein einzelner kompromittierter Server brächte bei homomorpher Verschlüsselung *gar keinen* Erkenntnisgewinn (ohne Schlüssel), bei MPC nur partielle Daten (ein Share ist ohne die anderen nichtssagend), bei Blockchain ebenfalls keinen Klartext (wenn richtig verschlüsselt). Gegenüber internen neugierigen Administratoren bieten daher alle Ansätze einen guten Schutz. Bei den externen Angriffsvektoren (z.B. Brute-Force, Kryptoanalyse) hängt die Sicherheit von den verwendeten primitiven Verfahren ab – alle fußen auf gut untersuchten hardness Annahmen (RSA/EC-EIGamal bei HE, Shamir-Shares und OT bei MPC, Hash-Kryptographie bei Blockchain). Insgesamt kann man attestieren: In puncto Datenschutz sind Homomorphe Verschlüsselung und MPC ausgezeichnet geeignet, während Blockchain primär als *Ergänzung* dient, um Vertrauenswürdigkeit und Integrität zu erhöhen. Keine der Methoden schwächt die biometrische Sicherheit (also die Fähigkeit, zwischen

richtigen und falschen Fingerabdrücken zu unterscheiden) direkt – außer eventuell indirekt durch approximative Algorithmen bei HE, was uns zum nächsten Punkt führt.

- **Erkennungsgenauigkeit und biometrische Performance:** Hier zeigt sich ein differenziertes Bild. Homomorphe Verfahren mussten in der Vergangenheit oft auf vereinfachte Matching-Modelle zurückgreifen (z.B. Fingercode statt detaillierter Minutien) [24]. Dies führte zu deutlich schlechteren Fehlerkennziffern (hohe EER) [24]. Moderne FHE-Systeme könnten zwar prinzipiell auch komplexere Vergleiche durchführen, aber die extreme Laufzeit machte dies unpraktisch. MPC hingegen konnte in den gezeigten Arbeiten die vollständigen Minutiendaten auswerten und erreichte damit die gleiche Genauigkeit wie ein klartextbasiertes System [24]. Aus biometrischer Sicht ist das ideal: keine Einbußen bei FAR/FRR, sodass die Sicherheit des Systems (im Sinne von Fälschungsresistenz) gleich hoch bleibt. Blockchain hat auf die rein biometrische Genauigkeit keinen Einfluss, da es eher eine Management-Schicht ist; relevant ist höchstens, dass Verzögerungen oder Paketverluste in einem verteilten Netzwerk zu Abbrüchen führen könnten – was aber die Fehlererkennung nicht beeinträchtigt, sondern nur den Durchsatz. In Bezug auf Skalierbarkeit (große Datenbanken, 1:N Suche) sind alle Methoden gefordert. HE könnte theoretisch eine Suchanfrage gegen viele Datensätze parallelisieren (z.B. mit Cloud-Hardware), leidet aber unter der Rechnerlast. MPC skaliert schlecht mit Anzahl Vergleichspaare, da Kommunikation quadratisch ansteigen kann. Blockchain skaliert je nach Protokoll begrenzt, vor allem was Transaktionsrate betrifft. Für den praktischen Einsatz großer AFIS-Systeme (Automated Fingerprint Identification Systems mit Millionen Datensätzen) müssten daher bei allen Ansätzen weitere architektonische Lösungen hinzukommen, etwa eine *Vorfilterung* der Kandidaten (Klassifizierung im verschlüsselten Raum) oder ein mehrstufiges Verfahren, um die Last zu senken [24]. Hier sind hybride Systeme denkbar: z.B. könnte man zuerst einen groben Homomorph-Abgleich mit einem schnell berechenbaren Merkmal machen, um Kandidaten einzugrenzen, und dann einen präzisen MPC-Vergleich der besten Treffer durchführen. Zusammengefasst: MPC bietet derzeit die beste Gewähr, die volle Genauigkeit eines Fingerabdruckverfahrens zu erhalten, während HE noch Genauigkeitsverluste in Kauf nimmt (aber perspektivisch aufholfähig ist). Blockchain hat keine direkte Auswirkung außer der möglichen Integration zusätzlicher Sicherheitsmaßnahmen (die aber ebenfalls keinen Genauigkeitsverlust bedingen sollten).
- **Leistung und Ressourcen:** Die Performance ist die Achillesferse aller Datenschutzlösungen. Ein normaler Fingerabdruckvergleich dauert in Software wenige Millisekunden. Homomorphe Vergleiche dauerten, wie gezeigt, in aktuellen Implementierungen über zwei Minuten pro Match* [27] – ein Faktor 10^4 langsamer. MPC konnte diese Lücke auf etwa einen Faktor 10–100 verkleinern (einige Sekunden) [24]. Das ist ein riesiger Fortschritt, aber für einige Anwendungen (z.B. Entsperren des Smartphones per Fingerabdruck, wo der Nutzer <1s erwartet) noch zu langsam. Dennoch: Für viele E-Government oder Cloud-Anwendungen könnten wenige Sekunden tolerabel sein, insbesondere wenn man bedenkt, dass Netzwerk-Latenzen oft dominieren. Blockchain-Einsatz bringt typischerweise ebenfalls sekunden- bis minutenlange Verzögerungen durch Konsens (z.B. 10 Sekunden bis ein Block mit der neuen Transaktion gefestigt ist, je nach Blockchain). Das ist problematisch für schnelle Authentifizierung – Lösungen könnten Off-Chain-Protokolle (State Channels) sein oder permissioned Blockchains mit schnelleren Abschlusszeiten. Auch der Ressourcenbedarf ist wichtig: HE erfordert enorme Rechenleistung (vor allem CPU; FHE ist schwer zu parallelisieren auf GPU, da sehr große Bitweiten und polynomiale Operationen). MPC benötigt hohe Netzwerkbandbreite und geringe Latenz zwischen den Servern, damit die vielen Kommunikationsrunden flott durchlaufen. Außerdem muss jeder MPC-Server ausreichend RAM/CPU haben, um die umfangreichen geheimen Datenstrukturen (z.B. riesige Boolesche Schaltkreise) zu handhaben. Blockchain-Knoten benötigen Speicher und Rechenleistung zum Validieren der Kette, was aber im Vergleich moderat ist – es skaliert mit der Anzahl Transaktionen, nicht direkt mit der Biometriedaten-Größe (weil man ja nur Hashes speichert). Insgesamt ist MPC hier der ressourcenintensivste Ansatz (verteilt auf mehrere Server), gefolgt von Homomorphie (zentral, aber rechnerisch schwer) und Blockchain (weniger rechenintensiv, aber verteilt redundant).
- **Komplexität und Reifegrad:** In puncto Implementierungskomplexität liegen ebenfalls Unterschiede vor. Homomorphe Verschlüsselung lässt sich – sofern man auf bestehende Bibliotheken zurückgreift – vergleichsweise *einfach* in ein Client-Server-Modell integrieren. Man verschlüsselt die Templates und schickt sie an den Server; das ist konzeptionell dem traditionellen Ansatz am nächsten. Die Schwierigkeit liegt hauptsächlich in der Parameterauswahl und Optimierung der kryptographischen Library, weniger im Systemdesign. MPC hingegen verlangt ein gründliches Verständnis des verteilten Protokolls und der sicheren Aufteilung. Die gesamte Applikationslogik muss in MPC-Komponenten (z.B. als arithmetischer Schaltkreis oder in einer domänenspezifischen MPC-Sprache) neu umgesetzt werden. Dies erfordert viel Aufwand und Testing – wie Mader *et al.* zeigten, mussten mehrere

algorithmische Änderungen vorgenommen werden, um z.B. Redundanzen zu entfernen und lineare anstelle von verzweigten Schleifen zu verwenden [24]. Das Fehlerrisiko (Bugs im Protokoll) ist nicht zu vernachlässigen, da die Debugging-Möglichkeiten begrenzt sind (man sieht ja nicht direkt, welche Zwischenwerte falsch laufen, ohne die Privatsphäre zu verletzen). Blockchain-Integrationen sind wiederum komplex auf Architekturebene: Man muss klassische IT-Systeme (Datenbanken, Matching-Server) mit einer dezentralen Ledger-Logik verbinden. Außerdem kommen hier neue Failure Modes ins Spiel – z.B. was passiert, wenn sich die Chain aufspaltet (Fork) oder ein Smart Contract einen Fehler hat? Diese Szenarien müssen mitbedacht werden. Vom *Reifegrad* her sind homomorphe Verschlüsselungsbibliotheken (Microsoft SEAL, PALISADE etc.) heute gut dokumentiert und anwendbar, es fehlt „nur“ an Geschwindigkeit. MPC-Frameworks wie MP-SPDZ, Sharemind, FRESCO haben ebenfalls einen Stand erreicht, der erste Realwelt-Projekte erlaubt, wie unsere Fallstudien zeigen. Blockchain im Bereich Digital Identity ist noch experimentell – es gibt Pilotprojekte (z.B. sovryn, EU ESSIF), aber einen allgemein anerkannten Standard, wie genau biometrische Authentifizierung über Blockchain laufen sollte, gibt es nicht. Hier spielt auch Regulierung mit hinein (DSGVO, E-IDAS-Verordnung etc.), die diese Ansätze teilweise erst noch einrahmen muss.

Zusammenfassend lässt sich die Eignung wie folgt einschätzen: Homomorphe Verschlüsselung ist theoretisch sehr attraktiv (einfaches Modell, starke Sicherheit), in der Praxis für vollwertiges Fingerabdruckmatching aber aktuell zu langsam. Secure MPC hat sich als praxisnäher erwiesen – es ermöglicht hochgenaue Vergleiche mit vertretbarem Mehraufwand und hat in prototypischen Umgebungen bereits den Härtestest bestanden [24]. Für Anwendungen, die mehrere vertrauenswürdige Parteien zulassen, ist MPC derzeit die beste Wahl für anonymes Matching. Distributed Ledgers schließlich stellen eine gute Ergänzung dar, wenn es um *Vertrauensmanagement, Dezentralisierung und Nachvollziehbarkeit* geht. Sie lösen nicht das Kernproblem des vertraulichen Vergleichs, können aber ein System bauen, in dem kein Teilnehmer allein alle Daten besitzt – was dem Ideal der *Datenminimierung* entspricht. In einem ganzheitlichen System könnten z.B. Blockchain und MPC kombiniert werden: Die Blockchain verteilt und sichert die biometrischen Referenzwerte, während MPC zwischen den Blockchain-Knoten den Abgleich durchführt. Solche Ansätze könnten in Zukunft besonders für grenzüberschreitende Identitätsprüfungen interessant sein, wo mehrere Behörden einander nicht voll vertrauen, aber dennoch kooperieren müssen (ähnlich dem No-Fly-List-Beispiel).

3.5 Fazit

Anonymes biometrisches Matching – insbesondere die privatsphärensichernde Verarbeitung von Fingerabdrücken – hat in den letzten Jahren enorme Fortschritte gemacht. Während früher ein Zielkonflikt zwischen *Sicherheit* und *Privatsphäre* bestand, zeigen moderne Kryptographie und verteilte Systeme Wege auf, diesen Konflikt weitgehend aufzulösen. Homomorphe Verschlüsselung und Secure Multi-Party Computation ermöglichen es, einen Fingerabdruckvergleich durchzuführen, ohne dass die beteiligten Instanzen die Fingerabdrücke selbst kennen. Die Machbarkeit wurde durch Forschungsprojekte eindrücklich belegt: So konnte erstmals ein minutenbasierter Abgleich unter MPC in wenigen Sekunden realisiert werden [24] – ein wichtiger Meilenstein. Distributed Ledgers wiederum eröffnen neue Architekturen, in denen biometrische Identitäten dezentral verwaltet und überprüft werden können, was das Vertrauen in solche Systeme erhöht.

Trotz aller Fortschritte stehen noch Herausforderungen an. Die Performance muss für manche Anwendungen weiter verbessert werden – insbesondere vollhomomorphe Verfahren benötigen hier weitere Innovationen, sei es durch algorithmische Verbesserungen oder leistungsfähigere Hardware (Stichwort Quantum Computing oder ASICs für FHE). Auch die Skalierung auf große Datenbanken und 1:N Matching erfordert zusätzliche Ansätze, etwa gestufte Verfahren oder mehrstufige Filter, damit die Verarbeitungszeit linear (und nicht exponentiell) mit der Datenbankgröße wächst. Ferner sind Standardisierungen notwendig: Derzeit existieren verschiedene Protokolle und Frameworks, aber es gibt keine einheitliche „Plug-and-Play“-Lösung. Für einen industriellen Einsatz müssten interoperable Standards geschaffen werden, die z.B. festlegen, wie ein verschlüsselter Fingerabdruckdatensatz formatiert ist, damit unterschiedliche Systeme zusammenspielen können.

Nicht zuletzt ist die rechtliche und gesellschaftliche Akzeptanz zu betrachten. Systeme für anonymes biometrisches Matching adressieren direkt Datenschutzbedenken – dies ist positiv und kann die Akzeptanz biometrischer Verfahren erhöhen. Dennoch muss transparent kommuniziert werden, wie die Technik funktioniert, um Vertrauen bei Nutzern aufzubauen. Da die Methoden komplex sind, besteht Aufklärungsbedarf, damit Entscheidungsträger verstehen, dass z.B. ein Abgleich „in der Cloud“ nicht zwangsläufig einen Datenschutzverstoß darstellt, solange fortgeschrittene kryptographische Verfahren im Einsatz sind.

Insgesamt zeigt dieser Bericht, dass anonymes biometrisches Matching kein Widerspruch in sich mehr ist, sondern ein erreichbares Ziel. Durch eine kluge Kombination aus Kryptographie (HE, MPC) und Systemdesign

(verteilte Ledger, sichere Hardware) können Fingerabdrücke verifiziert werden, ohne identifiziert zu werden – d.h. die Privatsphäre bleibt gewahrt. Für wissenschaftliche Leser ohne spezielle Vorkenntnisse mag es erstaunlich sein, dass derartiges möglich ist, doch die zitierten Arbeiten belegen eindrucksvoll das Potenzial. In naher Zukunft könnten entsprechende Verfahren in realen Anwendungen auftauchen, z.B. beim internationalen Identitätsmanagement oder beim sicheren Login über biometrische Merkmale, die *nur der Benutzer selbst* entschlüsseln kann. Der Weg von der Forschung zum Produkt ist zwar anspruchsvoll, aber angesichts der schnellen Fortschritte durchaus realistisch. Anonymes biometrisches Matching ist ein exemplarisches Feld, in dem Datenschutz durch Technik (Privacy by Design) verwirklicht wird – ein Ansatz, der in unserer datengesteuerten Welt zunehmend an Bedeutung gewinnt. Damit leistet die Forschung auf diesem Gebiet einen wichtigen Beitrag, biometrische Sicherheit und informationelle Selbstbestimmung in Einklang zu bringen.

4 RISK ASSESSMENT TOOLKIT UND VISION

4.1 Einleitung

Wie schon im Kapitel 3.3 (KI Methode Proof-of-Concept, AP2 / Task 2.2) dargestellt wird vom BMI als Ansatz für das Risk Assessment Framework die Erstellung von Abfragen mittels natürlicher Sprache als sehr sinnvoll einsetzbar erachtet.

Als technische Umsetzung wurde ein Agenten System gewählt, dass es ermöglicht anhand von verschiedenen Schritten zielgerichtet Ergebnisse von anderen Agenten zu ermitteln. Zum einen wird das LLM durch spezifizierte Prompts auf die Aufgabe vorbereitet (die verwendeten Datenbank Schemata werden mit eingegeben). Wenn vom LLM eine Abfrage erstellt wurde, wird diese dann von einem anderen Agenten genutzt um diese auf die SQL Datenbank anzuwenden. Zudem wird dem Operator (also dem PIU Mitarbeiter) immer die Möglichkeit gegeben, die SQL Abfragen zu kontrollieren und damit zu überprüfen, ob diese vernünftig umgesetzt wurde.

Eine wichtige Voraussetzung für die KI-Methode zur Erstellung der Abfragen mittels natürlicher Sprache über einen Chat-Bot liegt in darin, dass dieser Chat-Bot auf Systemen läuft, die von der IT-Abteilung des BMI bereitgestellt werden, da ansonsten die Datenschutzbestimmungen nicht erfüllt werden können. In einem System, dass in Echtbetrieb geht, darf z.B. Chat-GPT (von OpenAI) für so eine Aufgabe nicht verwendet werden, da dann Daten z.B. über den Aufbau der Datenbank einer nicht kontrollierbaren dritte Partei offenbart werden. Aus diesem Grund kommen für einen Echtbetrieb nur Systeme in Frage, die auf einem lokalen Server lauffähig sind.

Im Rahmen der Proof-of-Concept Implementierung wurde allerdings auf Chat-GPT 3.5 zurückgegriffen, da im Rahmen dieses Projekts die finanziellen Ressourcen für die Entwicklung eines Offlinesystems nicht vorhanden sind. Zudem wird im Rahmen dieses Projekts selbstverständlich nicht auf Datenbanken der PIU zugegriffen. AIT hat dazu eine eigene Datenbank bereitgestellt, die mit synthetisch generierten Daten angefüllt wurde. Der damit erstellte PoC wurde von Seiten der PIU Mitarbeiter als sehr positiv aufgenommen.

Als User-Interface wurde eine Chat-Bot Implementierung gewählt, da diese sehr intuitiv und schnell zu verstehen ist. Des Weiteren sind Ideen angedacht, wie Daten in Form von Netzwerken dargestellt werden können, diese Schritte sind aber noch nicht abschließend fertiggestellt.

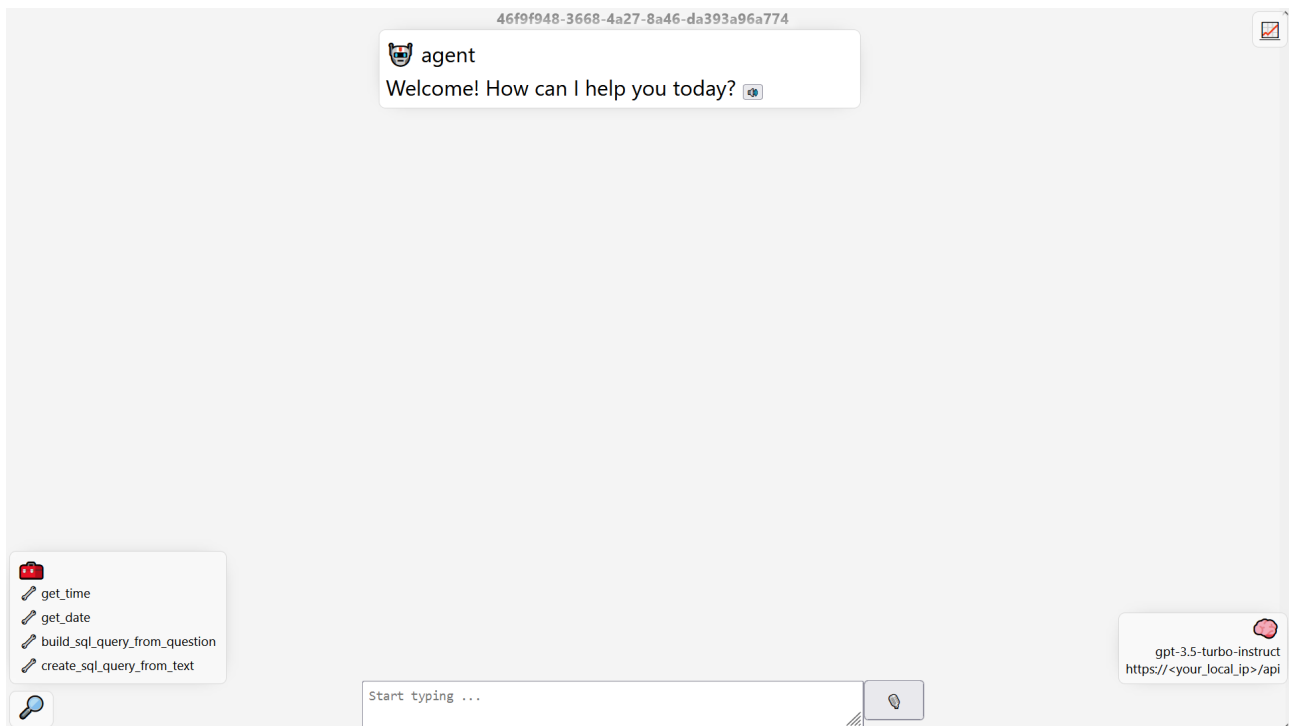
4.2 Proof-Of-Concept und User Interface Vision

In diesem Kapitel möchten wir als mittels einfacher Abbildungen eine Vorstellung des PoC-Systems geben. Es ist ein wenig gemixt mit dem nächsten Kapitel bzgl. der User Interface Vision, da hier ja diese Vision schon mit dargestellt wurde.

Gestartet wird das System über ein Webinterface, natürlich muss man sich noch dazu die entsprechenden Login Seiten vorstellen, die sicherstellen, dass nur berechtigte Mitarbeiter das System verwenden können. Diese sind hier nicht dargestellt, da dies in den PIU Abteilungen selbstverständlich eine Voraussetzung für jeglichen Einsatz von Abfrage Systeme gilt.

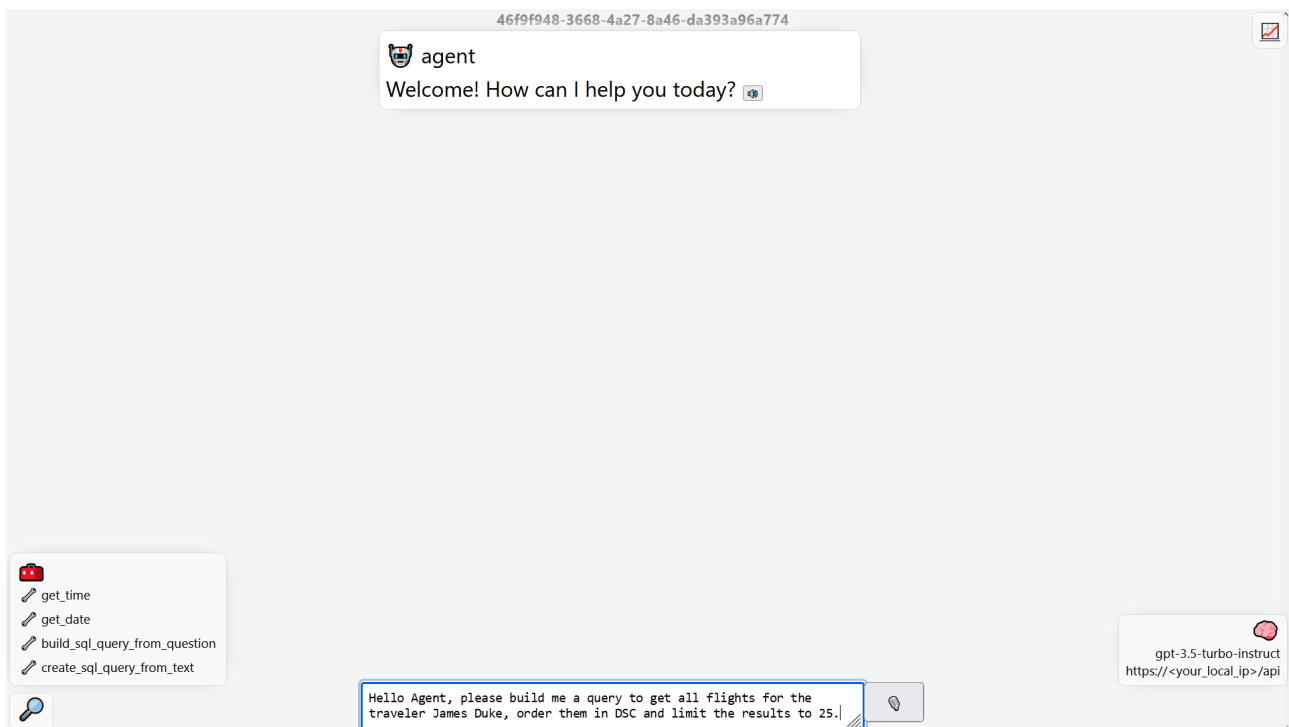
Das Interface ist als ChatBot umgesetzt, so dass es in einem natürlich Sprachlichen Prozess gut einfügt. Als Sprache ist Englisch vorgesehen, da sich so die Erstellung für internationale Präsentation der Aufwand deutlich verringert.

In der folgenden Abbildung ist die erste Website, die der Nutzer nach dem Einloggen gezeigt bekommt dargestellt:

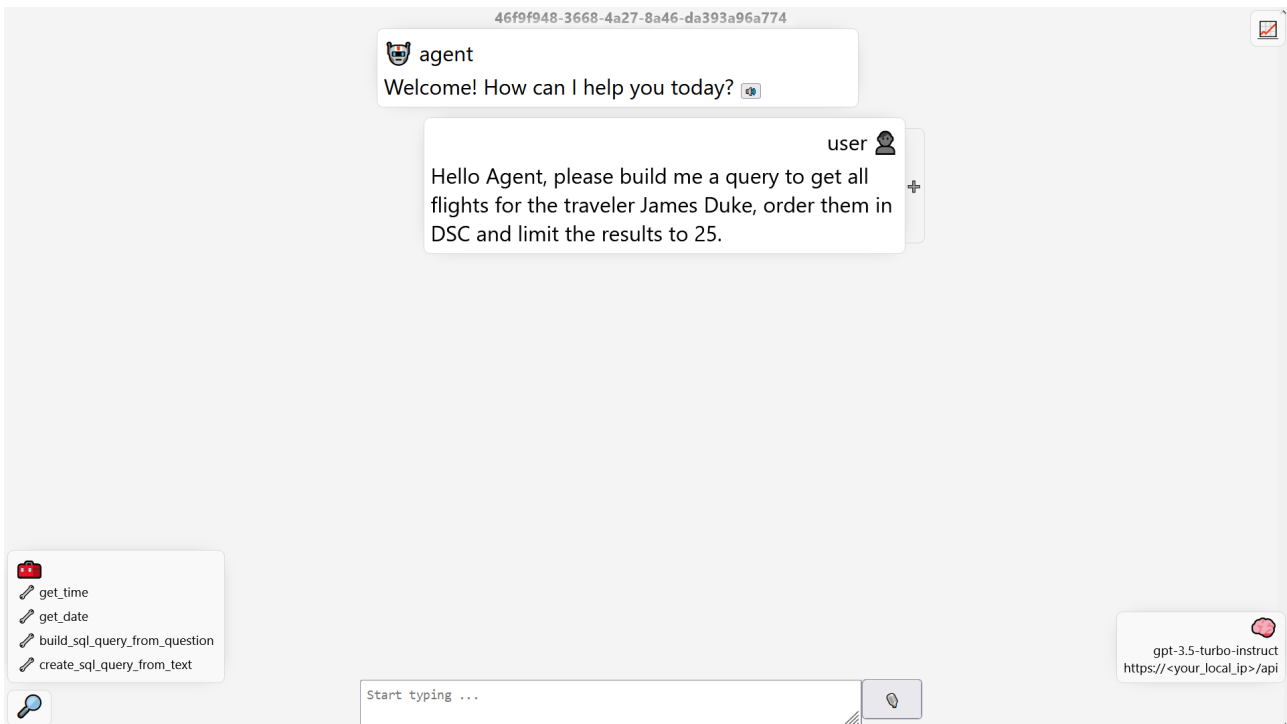


Er wird entsprechend begrüßt und darauf hingewiesen, dass er jetzt Fragen stellen kann. Unten auf der Website sieht man die Eingabebox, in die eine erste Frage eingegeben werden kann. Auf der rechten sieht man unten das verwendete LLM Modell, in diesem Fall wie schon erwähnt das gpt-3.5-turbo-instruct Modell. Hier werden in Zukunft mehrere lokale Modelle zur Auswahl stehen, so dass man die praktische Einsatzfähigkeit verschiedener LLM Modelle untersuchen kann. Auf der linken Seite werden die verschiedenen Agenten angezeigt, die in einer zukünftigen Version entsprechend erweitert werden.

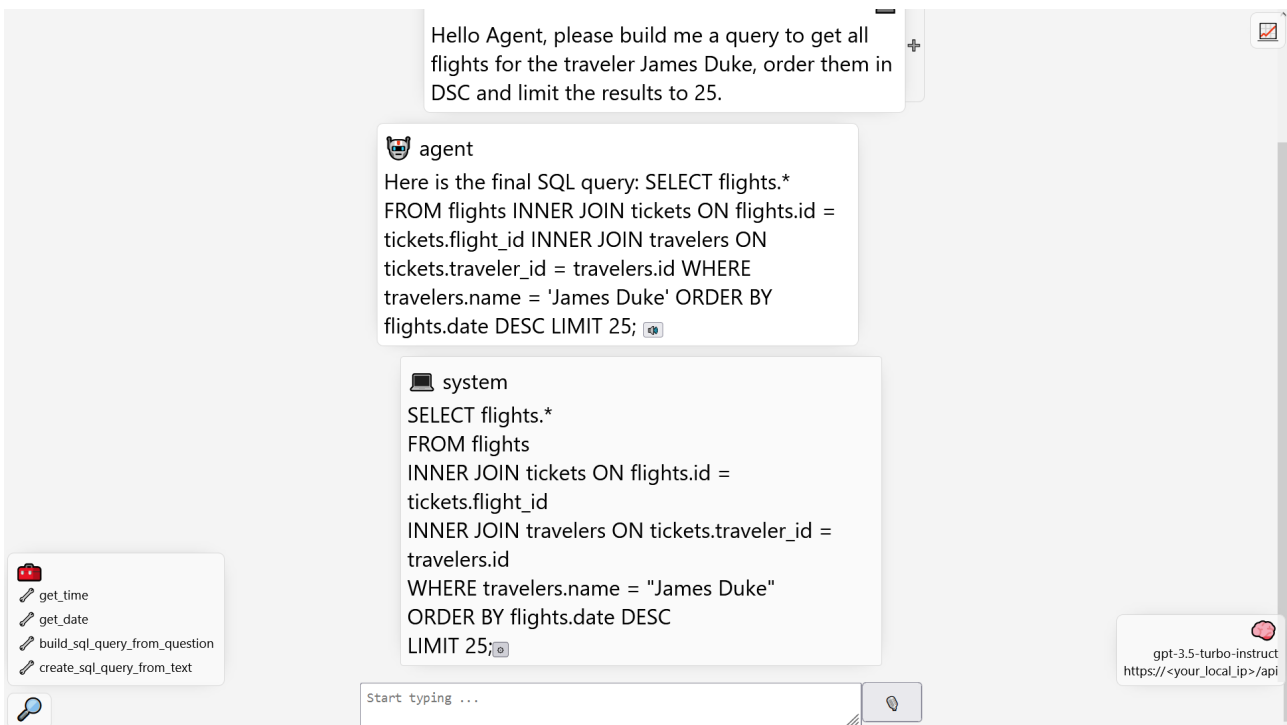
Der User gibt nun die Frage seiner Anfrage in das untere Eingabefeld ein:



Sobald die Eingabetaste (return) gedrückt wird, wird die ChatBot Historie angezeigt, d.h. der Eingabetext rückt nach oben und das Eingabefeld unten wird wieder geleert und steht für neue Eingaben bereit:



Durch die Ausrichtung der Eingabebboxen im oberen Bereich wird angezeigt, wo System Ausgaben/Fragen stehen (nach links verschobene Box), und wo die Historie der Usereingaben sind (nach rechts verschobene Box). Nach wenigen Sekunden fügt der Agent die gefundene Antwort hinzu. Zudem erscheint eine weitere Box, die aus der Antwort den SQL Teil herausfiltert:



Diese vom Agenten System herausgefilterte SQL Anfrage wird dann entsprechend an den SQL Agent weitergereicht, der die Abfrage auf der SQL Datenbank ausführt und die Ergebnisse am rechten Rand des Bildschirms darstellt:

Hello Agent, please build me a query to get all flights for the traveler James Duke, order them in DSC and limit the results to 25.

agent

Here is the final SQL query: SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = 'James Duke' ORDER BY flights.date DESC LIMIT 25;

system

SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = "James Duke" ORDER BY flights.date DESC LIMIT 25;

get_time
get_date
build_sql_query_from_question
create_sql_query_from_text

Start typing ...

id	origin	destination	date
79	South Taylor	New Bradleyton	2025-06-06
35	Sloantown	Port Cherylfort	2025-04-27
46	Heatherside	Vienna	2025-04-22
51	Christinaside	Parsonsville	2025-02-14
16	Snyderville	Williamsonshire	2024-12-25
16	Snyderville	Williamsonshire	2024-12-25
67	West Michele	North Hannahport	2024-12-25
39	Vienna	Port Tony	2024-12-01
36	North Joshua	Vienna	2024-11-08
3	Richardchester	New Maxside	2024-10-18
91	New Melissa	Port Courtneychester	2024-09-29
4	Vienna	Port Ruth	2024-08-23
72	Emilyfurt	Johnsonshire	2024-08-15
69	New Christopherport	South Feliciamouth	2024-08-10
9	Bairdstad	Vienna	2024-08-08
53	Christopherside	East Coreyborough	2024-07-23
40	Greenhaven	South Richard	2024-06-22

Query:
SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = "James Duke" ORDER BY flights.date DESC LIMIT 25;

gpt-3.5-turbo-instruct
https://<your_local_ip>/api

Da die Anzahl der Resultate recht hoch ist, kann man versuchen, die Anfrage zu verfeinern, z.B. derart, dass nur Flüge nach Wien betrachtet werden sollen, das wird entsprechend wieder in das Eingabefeld unten eingtragen:

Hello Agent, please build me a query to get all flights for the traveler James Duke, order them in DSC and limit the results to 25.

agent

Here is the final SQL query: SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = 'James Duke' ORDER BY flights.date DESC LIMIT 25;

system

SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = "James Duke" ORDER BY flights.date DESC LIMIT 25;

get_time
get_date
build_sql_query_from_question
create_sql_query_from_text

Thank you.
Build a refined query only considering flights with a destination Vienna.

id	origin	destination	date
79	South Taylor	New Bradleyton	2025-06-06
35	Sloantown	Port Cherylfort	2025-04-27
46	Heatherside	Vienna	2025-04-22
51	Christinaside	Parsonsville	2025-02-14
16	Snyderville	Williamsonshire	2024-12-25
16	Snyderville	Williamsonshire	2024-12-25
67	West Michele	North Hannahport	2024-12-25
39	Vienna	Port Tony	2024-12-01
36	North Joshua	Vienna	2024-11-08
3	Richardchester	New Maxside	2024-10-18
91	New Melissa	Port Courtneychester	2024-09-29
4	Vienna	Port Ruth	2024-08-23
72	Emilyfurt	Johnsonshire	2024-08-15
69	New Christopherport	South Feliciamouth	2024-08-10
9	Bairdstad	Vienna	2024-08-08
53	Christopherside	East Coreyborough	2024-07-23
40	Greenhaven	South Richard	2024-06-22

Query:
SELECT flights.* FROM flights INNER JOIN tickets ON flights.id = tickets.flight_id INNER JOIN travelers ON tickets.traveler_id = travelers.id WHERE travelers.name = "James Duke" ORDER BY flights.date DESC LIMIT 25;

gpt-3.5-turbo-instruct
https://<your_local_ip>/api

VERSIONSNUMMER |

29

und durch Eingabebestätigung (Return Taste) an das System weitergegeben. Die Anfrage wird entsprechend verfeinert, das SQL Query und auch die Abfrageergebnisse entsprechend aktualisiert. Zu beachten, ist dabei der Effekt auf die Anzahl der Ergebnisse (in der Box im oberen rechten Bereich des Bildschirms):

The screenshot shows a chat interface with three participants: user, agent, and system. The user asks for flights to Vienna. The agent responds that there are no other valid interpretations. The system generates a SQL query. On the right, a table shows flight results.

user

Thank you.
Build a refined query only considering flights with a destinaion Vienna.

agent

There are no other valid interpretations for the question.

system

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = 'James Duke' AND
flights.destination = 'Vienna'
ORDER BY flights.date DESC
LIMIT 25;
```

Query:

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = 'James Duke' AND
flights.destination = 'Vienna'
ORDER BY flights.date DESC
LIMIT 25;
```

id	origin	destination	date
46	Heatherside	Vienna	2025-04-22
36	North Joshua	Vienna	2024-11-08
9	Bairdstad	Vienna	2024-08-08

get_time
get_date
build_sql_query_from_question
create_sql_query_from_text

Start typing ...

gpt-3.5-turbo-instruct
https://<your_local_ip>/api

Wenn diese Ergebnisse aber nicht zufriedenstellend sind, kann man auch einfach in der Abfrage zurück gehen, und durch drücken des +-Zeichens neben der letzten Eingabe eine Verzweigung erstellen:

The screenshot is identical to the one above, but with a red arrow pointing to the plus sign next to the user's input. A red box with the word 'Click' is placed next to the arrow.

user

Thank you.
Build a refined query only considering flights with a destinaion Vienna.

agent

There are no other valid interpretations for the question.

system

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = 'James Duke' AND
flights.destination = 'Vienna'
ORDER BY flights.date DESC
LIMIT 25;
```

Query:

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = 'James Duke' AND
flights.destination = 'Vienna'
ORDER BY flights.date DESC
LIMIT 25;
```

id	origin	destination	date
46	Heatherside	Vienna	2025-04-22
36	North Joshua	Vienna	2024-11-08
9	Bairdstad	Vienna	2024-08-08

get_time
get_date
build_sql_query_from_question
create_sql_query_from_text

Thank you.
Build a refined query only considering flights between 2024-07-31 and 2025-01-01.

Click

gpt-3.5-turbo-instruct
https://<your_local_ip>/api

Unter der letzten Abfrage werden dann die verschiedenen Zweige der Abfragen dargestellt, die gerade aktuell verfolgte Abfrage wird dabei hervorgehoben

The screenshot shows a SQL query editor interface. On the left, a query tree displays the structure of the query, with the current query highlighted. The query is a SELECT statement that joins the flights, tickets, and travelers tables, filtering for flights by James Duke and ordering by date. The main area shows the SQL query text. On the right, a results table displays the query output, showing columns for id, origin, destination, and date. The table contains 20 rows of data. Below the query tree, there are two buttons: 'get_time' and 'get_date'. Below the results table, there are two buttons: 'build_sql_query_from_question' and 'create_sql_query_from_text'. At the bottom, there is a search bar with the text 'Start typing ...' and a 'user' profile icon.

```
tickets.flight_id INNER JOIN travelers ON
tickets.traveler_id = travelers.id WHERE
travelers.name = 'James Duke' ORDER BY
flights.date DESC LIMIT 25;
```

```
system
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id =
tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id =
travelers.id
WHERE travelers.name = "James Duke"
ORDER BY flights.date DESC
LIMIT 25;
```

Thank you. Build a refined query only considering flights with a destination Vienna.

Thank you. Build a refined query only considering flights between 2024-07-31 and 2025-01-01.

Thank you. Build a refined query only considering flights between 2024-07-31 and 2025-01-01.

Thank you.

Build a refined query only considering flights between 2024-07-31 and 2025-01-01.

Start typing ...

id	origin	destination	date
79	South Taylor	New Bradleyton	2025-06-06
35	Sloantown	Port Cherylfort	2025-04-27
46	Heatherside	Vienna	2025-04-22
51	Christinaside	Parsonsville	2025-02-14
16	Snyderville	Williamsonshire	2024-12-25
16	Snyderville	Williamsonshire	2024-12-25
67	West Michele	North Hannahport	2024-12-25
39	Vienna	Port Tony	2024-12-01
36	North Joshua	Vienna	2024-11-08
3	Richardchester	New Maxside	2024-10-18
91	New Melissa	Port Courtneychester	2024-09-29
4	Vienna	Port Ruth	2024-08-23
72	Emilyfurt	Johnsonshire	2024-08-15
69	New Christopherport	South Feliciamouth	2024-08-10
9	Bairdstad	Vienna	2024-08-08
53	Christopherside	East Coreyborough	2024-07-23
40	Greenhaven	South Richard	2024-06-22

Query:

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = "James Duke"
ORDER BY flights.date DESC
LIMIT 25;
```

gpt-3.5-turbo-instruct
https://<your_local_ip>/api

Und anschließend die Eingabe interpretiert, die neue SQL Query ermittelt und die Ergebnisse aktualisiert

The screenshot shows the same SQL query editor interface as before, but with the query refined to include a date range. The query tree on the left shows the updated query structure. The main area displays the new SQL query text. The results table on the right shows the updated query output, which now only includes flights between 2024-07-31 and 2025-01-01. The table contains 5 rows of data. The buttons and search bar at the bottom remain the same.

```
agent
The final SQL query is: SELECT flights.* FROM
flights INNER JOIN tickets ON flights.id =
tickets.flight_id INNER JOIN travelers ON
tickets.traveler_id = travelers.id WHERE
travelers.name = 'James Duke' AND flights.date
BETWEEN '2024-07-31' AND '2025-01-01' ORDER
BY flights.date DESC LIMIT 25;
```

```
system
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id =
tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id =
travelers.id
WHERE travelers.name = 'James Duke'
AND flights.date BETWEEN '2024-07-31' AND
'2025-01-01'
ORDER BY flights.date DESC
LIMIT 5;
```

Start typing ...

id	origin	destination	date
16	Snyderville	Williamsonshire	2024-12-25
16	Snyderville	Williamsonshire	2024-12-25
67	West Michele	North Hannahport	2024-12-25
39	Vienna	Port Tony	2024-12-01
36	North Joshua	Vienna	2024-11-08

Query:

```
SELECT flights.*
FROM flights
INNER JOIN tickets ON flights.id = tickets.flight_id
INNER JOIN travelers ON tickets.traveler_id = travelers.id
WHERE travelers.name = 'James Duke'
AND flights.date BETWEEN '2024-07-31' AND '2025-01-01'
ORDER BY flights.date DESC
LIMIT 5;
```

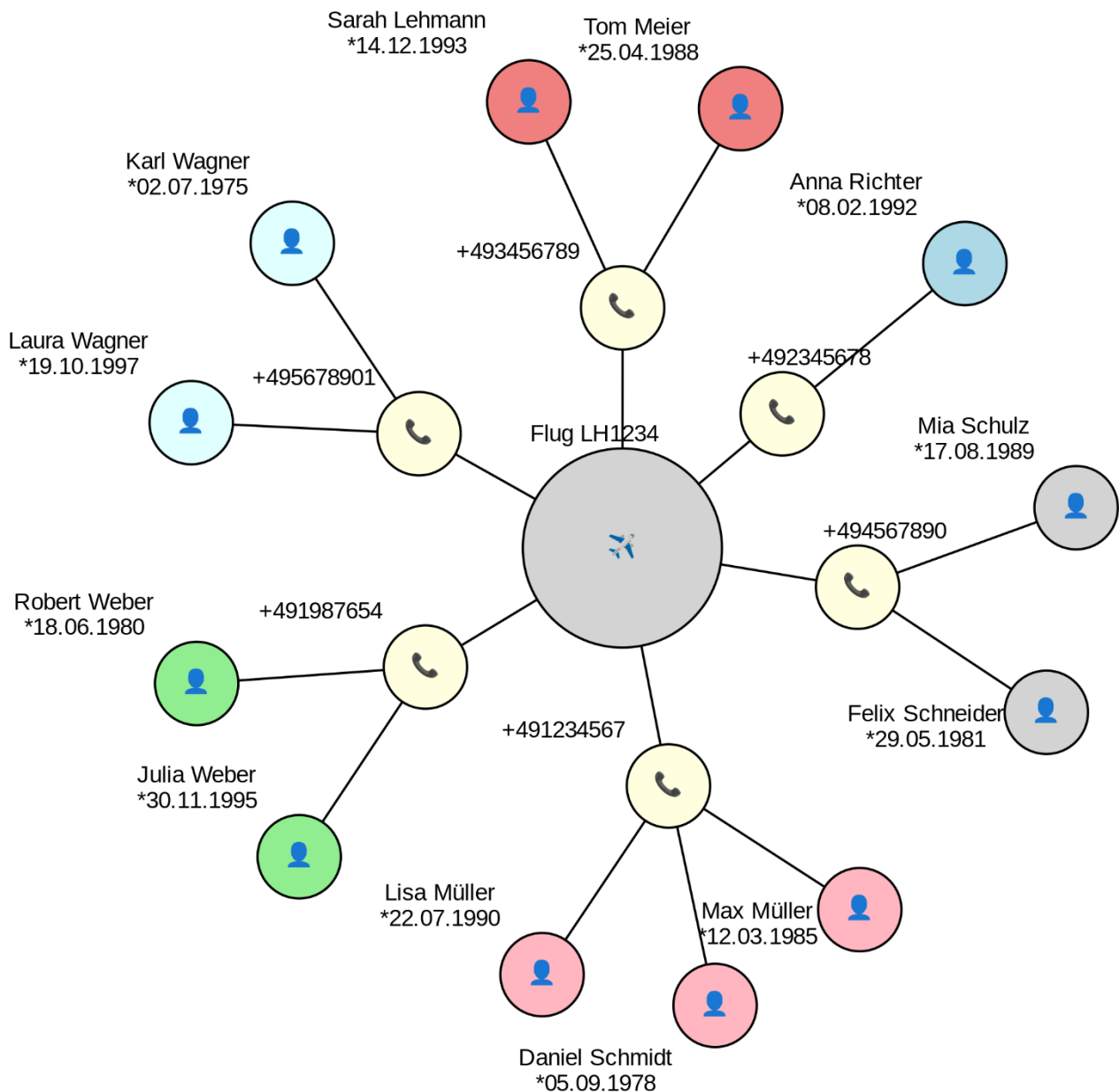
gpt-3.5-turbo-instruct
https://<your_local_ip>/api

Natürlich kann man jetzt die Abfrage weiter verfeinern, bis man mit dem Ergebnis zufrieden ist.

Durch diese Art der Abfrageerstellung kann man den ganzen Prozess deutlich einfacher gestaltet, ermöglicht einen schnellen Änderungsprozess, bei dem einfach auf frühere Ergebnisse zurückgesprungen werden können.

Die Vorstellung den technischen Prozess der Abfrageerstellung derart zu vereinfachen, wurde bei den Mitarbeitern der PIU sehr positiv aufgenommen. Es verringert auch die logische Komplexität, wenn man die SQL Abfrage direkt in der SQL Abfragesprache formulieren möchte. Mit so einen Ansatz wird auch die Einarbeitung neuer Mitarbeiter in der PIU deutlich vereinfacht.

Eine weitere Vision beim User Interface liegt insbesondere in der Darstellung von Zusammenhängen zwischen Telefonnummer, die für die Buchung verwendet werden und Fluggästen. Zum Beispiel können so für einem bestimmten Flug alle Passagiere dargestellt werden und wie diese mit den Telefonnummern verbunden sind:



Das ist eine Analysefunktion, die auch verschiedene der PIU Tools anbieten. Kombiniert mit einem ChatBot können so direkt und Natürlich-Sprachig die Analysen dargestellt werden. Das Konzept des agentenbasierten Systems ermöglicht damit eine sehr einfache Nutzungsmöglichkeit. Das Wissen um die interne, verwendete Datenbankstruktur rückt in den Hintergrund, das Analysewissen um Zusammenhänge in den Vordergrund. Gerade in Hinblick darauf, dass häufig Personen aus dem Polizeidienst in den PIUs arbeiten die beste Voraussetzung um deren, schon im Polizeidienst gesammelten Erfahrungen und Wissen bzgl. der Zusammenhänge im Kriminalitätsbereich, in der Analyse zum Erfolg zu führen.

5 DISSEMINATION AKTIVITÄTEN

Im Rahmen der Aufgabe 2.3 wurden folgende drei Veröffentlichungen erzielt:

- Masterarbeit Julia Mader. <https://publications.ait.ac.at/de/publications/privacy-preserving-fingerprint-matching>

- Mader, Julia, Lorünser, Thomas. "Feasibility of Privacy Preserving Minutiae-Based Fingerprint Matching: ICISSP 2024 : 10th International Conference on Information Systems Security and Privacy." In Proceedings of the 10th International Conference on Information Systems Security and Privacy, edited by Julia Mader, 1:899–906. Rome, Italy, 2024. <https://doi.org/10.5220/0012472300003648>.
- Julia Mader, Florian Wohner, Laurenz Ruzicka, and Thomas Loruenser. 2024. Towards Real-Time Privacy-Preserving Minutiae Matching. In Proceedings of the 23rd Workshop on Privacy in the Electronic Society (WPES '24), October 14–18, 2024, Salt Lake City, UT, USA. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3689943.3695049>

Des Weiteren hat Christian Bauer die neue natürlichsprachliche Suchen auf Chat-Bot basis auf folgenden internationalen Zusammenkünften vorgestellt:

- Informal Working Group Europe
- Working Group PNR der Europäischen Kommission
- UN-OCT Meeting

Für die Vorstellung der Methoden wurde ein entsprechendes kurzes Video erstellt, über dass die Arbeitsweise des ChatBot-Ansatzes illustriert wird.

6 LITERATUR

- [1] Bundesministerium für Inneres, "Fluggastdaten besser nutzen," *Öffentliche Sicherheit*, vol. 5–6, 2023, Available: https://www.bmi.gv.at/magazin/2023_05_06/15_Europaeische_Union.aspx
- [2] Heise online, "Rat stimmt zu: Neue EU-vorschriften zum sammeln von fluggastdaten stehen," Dec. 2024, Available: <https://www.heise.de/news/news/Rat-stimmt-zu-Neue-EU-Vorschriften-zum-Sammeln-von-Fluggastdaten-stehen-10197869.html>
- [3] Bundeskanzleramt, "Erläuterungen zum PNR-gesetz," 2018, Available: https://www.bundeskanzleramt.gv.at/dam/jcr:40077ab2-7ec7-4705-8fd9-8e463fb0cfd4/21_20_erlaeu.pdf
- [4] T. Rudl, "API-verordnung: EU weitert überwachung von flügen aus - mit abstrichen," Mar. 2024, Available: <https://netzpolitik.org/2024/api-verordnung-eu-weitet-ueberwachung-von-fluegen-aus-mit-abstrichen>
- [5] Parlament Österreich, "Bundesgesetz über die verarbeitung von fluggastdaten zur vorbeugung, verhinderung und aufklärung von terroristischen und bestimmten anderen straftaten (PNR-gesetz); bundeskriminalamt-gesetz, änderung (4/ME)," 2018, Available: <https://www.parlament.gv.at/gegenstand/XXVI/ME/4>
- [6] I.-Z. für Rechtsinformatik und Informationsrecht, "Urteil des VG wiesbaden zur fluggastdatenverarbeitung," May 2020, Available: <https://www.jurpc.de/jurpc/show?id=2020008>
- [7] K. Belgien, "FAQ zur PNR-verarbeitung," Jul. 2024, Available: <https://www.crisiscenter.be/de/was-tut-das-nationale-krisenzentrum/national-travel-targeting-center/pnr-verarbeitung-von>
- [8] Datenschutz Sachsen-Anhalt, "Artikel 8 - datenübermittlungspflichten der fluggesellschaften," Available: <https://datenschutz.sachsen-anhalt.de/recht/vorschriften/europarecht/richtlinie-eu-2016681-des-europaeischen-parlaments-und-des-rates-pnr-richtlinie/artikel-8-datenuebermittlungspflichten-der-fluggesellschaften>
- [9] LabourNet Germany, "Analyse zur PNR-auswertung in deutschland," Jun. 2024, Available: <https://labournet.de/interventionen/grundrechte/grundrechte-all/verfassungsschutz/auswertung-von-fluggastdaten-pnr/>
- [10] C. Kurz, "EuGH-urteil beschränkt massenüberwachung bei flugreisen." Netzpolitik.org, 16. Jan. 2024, Jun. 21, 2022. Available: <https://netzpolitik.org/2024/api-verordnung-eu-weitet-ueberwachung-von-fluegen-aus-mit-abstrichen>
- [11] B. für den Datenschutz, "Umsetzung des EuGH-urteils zur PNR-richtlinie," Dec. 2022, Available: https://www.bfdi.bund.de/SharedDocs/Pressemitteilungen/DE/2022/13_EDSA-PNR-Richtlinie.html
- [12] epicenter.works, "PNR fluggastdatenspeicherung." Webseite epicenter.works, abgerufen am 15.07.2019, Jul. 15, 2019. Available: <https://epicenter.works/thema/pnr>

- [13] “Critical technologies and social innovation schemes report.” [Online]. Available: <https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds=080166e5d7278606&appId=PPGMS>
- [14] R. Namazov, “Application of advance passenger information (API) and passenger name record (PNR) security systems by using travel information.” [Online]. Available: <https://border-security-report.com/wp-content/uploads/2022/07/API-PNR-Namazov-research.pdf>
- [15] “Second in-person international user community meeting on goTravel - news - portail de la police grand-ducale - luxembourg.” [Online]. Available: <https://police.public.lu/en/actualites/2024/10/semaine-43/ct-travel.html>
- [16] “HERMES: Screening solution.” [Online]. Available: <https://www.wcc-group.com/id-security/products/hermes-passenger-screening-solution/>
- [17] “WCC applies artificial intelligence to passenger screening | WCC group.” [Online]. Available: <https://www.wcc-group.com/company/post/2021/05/19/wcc-successfully-applies-artificial-intelligence-to-passenger-screening>
- [18] “SITA | SITA intelligence and targeting.” [Online]. Available: <https://sita.aero/solutions/sita-at-borders/border-management/sita-intelligence-and-targeting/>
- [19] “As travel surges, how do we drive down risk at the border?” [Online]. Available: <https://airportindustry-news.com/as-travel-surges-how-do-we-drive-down-risk-at-the-border/>
- [20] “SITA | API PNR gateway.” [Online]. Available: <https://sita.aero/solutions/sita-at-borders/border-management/sita-api-pnr-gateway/>
- [21] “<https://sita.aero/globalassets/docs/brochures/sita-intelligence-and-targeting-product-brochure.pdf>.” [Online]. Available: <https://sita.aero/globalassets/docs/brochures/sita-intelligence-and-targeting-product-brochure.pdf>
- [22] “Cybersecurity, biometrics, AI in investment focus, SITA report shows.” [Online]. Available: <https://aviationweek.com/aerospace/emerging-technologies/daily-memo-cybersecurity-biometrics-ai-investment-focus-sita-report>
- [23] “Travizory API-PNR targeting system - travizory.” [Online]. Available: <https://travizory.com/what-we-do/advance-passenger-information>
- [24] J. Mader and T. Lorünser, “Feasibility of privacy preserving minutiae-based fingerprint matching.” INSTICC; SciTePress, pp. 899–906, 2024. doi: [10.5220/0012472300003648](https://doi.org/10.5220/0012472300003648).
- [25] “File:fingerprints minutiae patterns representation.jpg - wikimedia commons.” [Online]. Available: https://commons.wikimedia.org/wiki/File:Fingerprints_Minutiae_Patterns_Representation.jpg
- [26] W. Yang, S. Wang, H. Cui, Z. Tang, and Y. Li, “A review of homomorphic encryption for privacy-preserving biometrics,” *Sensors*, vol. 23, no. 7, 2023, doi: [10.3390/s23073566](https://doi.org/10.3390/s23073566).
- [27] T. Kim, Y. Oh, and H. Kim, “Efficient privacy-preserving fingerprint-based authentication system using fully homomorphic encryption,” *Security and Communication Networks*, vol. 2020, no. 1, p. 4195852, 2020, doi: <https://doi.org/10.1155/2020/4195852>.
- [28] “2024: Blockchain & biometrics transform ID verification.” [Online]. Available: <https://www.rapidinnovation.io/post/the-future-of-identity-verification-blockchain-and-biometric-integration-in-2024>
- [29] S. Sharma and R. Dwivedi, “A survey on blockchain deployment for biometric systems,” *IET Blockchain*, vol. 4, no. 2, pp. 124–151, 2024, doi: <https://doi.org/10.1049/blc2.12063>.