



Bundesamt  
für Sicherheit in der  
Informationstechnik

## Studie:

# „Untersuchung der Leistungsfähigkeit von biometrischen Verifikationssystemen – BioP II“

## Öffentlicher Abschlussbericht



Bundesamt  
für Sicherheit in der  
Informationstechnik



Bundeskriminalamt

**secu**net

Version 2.0

23.08.2005



---

## Inhalt

<b>1</b>	<b>Vorwort.....</b>	<b>7</b>
<b>2</b>	<b>Management Summary .....</b>	<b>9</b>
<b>3</b>	<b>Einleitung .....</b>	<b>21</b>
3.1	Testziel.....	21
3.2	Systemauswahl.....	22
3.3	Projektphasen.....	23
3.4	Aufbau des Dokuments.....	23
<b>4</b>	<b>Beschreibung der verwendeten Referenzbasen .....</b>	<b>25</b>
4.1	Erstellung von Bilddateien als Referenzbasis.....	26
4.1.1	<i>Erstellung der Gesichtsbilder (RefID1).....</i>	<i>26</i>
4.1.2	<i>Erstellung der Fingerabdruckbilder (RefID3 und RefID5).....</i>	<i>30</i>
4.1.3	<i>Erstellung der Irisbilder (RefID7).....</i>	<i>32</i>
4.2	Erstellung von Templates als Referenzbasis.....	34
4.3	Bereitstellung der Referenzdaten für den Test.....	35
4.3.1	<i>Allgemeiner Aufbau der CBEFF-Struktur.....</i>	<i>35</i>
4.3.2	<i>Speicherung der primären und sekundären Merkmale für Finger und Iris.....</i>	<i>37</i>
<b>5</b>	<b>Beschreibung der Systeme .....</b>	<b>39</b>
5.1	Systemkomponenten.....	39
5.1.1	<i>GE-System von Cognitec .....</i>	<i>39</i>
5.1.2	<i>FE-System von Bundesdruckerei/NEC.....</i>	<i>40</i>
5.1.3	<i>FE-System von Dermalog.....</i>	<i>41</i>
5.1.4	<i>IE-System von SDI.....</i>	<i>42</i>
5.2	Funktionsweise der Systeme.....	44
5.2.1	<i>Enrolment-Modus.....</i>	<i>44</i>
5.2.2	<i>Verifikationsmodus.....</i>	<i>46</i>
5.2.3	<i>Batch-Modus.....</i>	<i>50</i>
<b>6</b>	<b>Feldtestbeschreibung.....</b>	<b>51</b>
6.1	Testbedingungen.....	51
6.1.1	<i>Testpopulation.....</i>	<i>51</i>
6.1.2	<i>Teilnehmermotivation.....</i>	<i>57</i>
6.1.3	<i>Datenschutz.....</i>	<i>58</i>
6.1.4	<i>Testumgebung.....</i>	<i>58</i>
6.1.5	<i>Architektur des Testaufbaus.....</i>	<i>60</i>

---

6.1.6	<i>Systemkonfiguration</i> .....	62
6.1.7	<i>Wartung und Pflege der biometrischen Systeme</i> .....	65
6.2	Testdurchführung .....	65
6.2.1	<i>Anwerbung und Anmeldung der Testpersonen</i> .....	65
6.2.2	<i>Feldtest</i> .....	81
<b>7</b>	<b>Feldtestauswertung</b> .....	<b>84</b>
7.1	Auswertungskonzept .....	84
7.1.1	<i>Vergleichstypen</i> .....	84
7.1.2	<i>Bewertung der Erkennungsleistungen</i> .....	84
7.2	Testergebnisse .....	90
7.2.1	<i>Definition der Basisdatenmengen</i> .....	90
7.2.2	<i>Failure-To-Enrol Rate (FTE)</i> .....	93
7.2.3	<i>Erkennungsleistungen</i> .....	95
7.2.4	<i>Generalisierte Erkennungsleistungen</i> .....	124
7.2.5	<i>Korrelationen</i> .....	126
7.2.6	<i>Zeitliche Performance</i> .....	131
7.2.7	<i>Template-Größe</i> .....	132
7.2.8	<i>Allgemeine Ergebnisse zu Systemen und Herstellern</i> .....	132
<b>8</b>	<b>Weiterführende Untersuchungen</b> .....	<b>134</b>
8.1	Benutzbarkeit und Akzeptanz .....	134
8.1.1	<i>Datenerhebung</i> .....	134
8.1.2	<i>Benutzbarkeit</i> .....	135
8.1.3	<i>Akzeptanz</i> .....	139
8.1.4	<i>Empfehlungen zur Gestaltung der Systeme</i> .....	139
8.2	Robustheit und Langlebigkeit der Sensoren .....	141
8.2.1	<i>GE-System von Cognitec</i> .....	142
8.2.2	<i>FE-System von Bundesdruckerei/NEC</i> .....	142
8.2.3	<i>FE-System von Dermalog</i> .....	144
8.2.4	<i>IE-System von SD Industries</i> .....	145
8.2.5	<i>Zusammenfassung</i> .....	146
8.3	Speicherung biometrischer Daten auf RF-Token .....	146
<b>9</b>	<b>Bewertungsschema</b> .....	<b>148</b>
9.1	Aufbau des Bewertungsschemas .....	148
9.2	Bewertungskriterien .....	148

---

9.2.1	<i>Erkennungsleistung/Kennzahlen</i> .....	148
9.2.2	<i>Systemergebnisse</i> .....	149
9.2.3	<i>Benutzerakzeptanz</i> .....	150
9.2.4	<i>Benutzbarkeit</i> .....	150
9.2.5	<i>Sicherheit</i> .....	151
9.2.6	<i>Herstellerunterstützung</i> .....	151
9.2.7	<i>Nicht zu bewertende Kriterien</i> .....	151
9.3	Klassifikation der Ergebnisse .....	152
9.4	Systembewertung.....	159
<b>10</b>	<b>Zusammenfassung und Interpretation der Ergebnisse</b> .....	<b>162</b>
10.1	Erkennungsleistung der biometrischen Verfahren .....	162
10.2	Einsetzbarkeit ICAO-Bilddateien als Referenz .....	165
10.3	Benutzbarkeit und Akzeptanz .....	166
10.3.1	<i>Benutzbarkeit</i> .....	166
10.3.2	<i>Akzeptanz</i> .....	166
10.4	Eignung der getesteten Verfahren.....	166
10.4.1	<i>Gesichtserkennung</i> .....	166
10.4.2	<i>Fingerabdruckerkennung</i> .....	167
10.4.3	<i>Iriserkennung</i> .....	167
10.4.4	<i>Vergleichende Bewertung</i> .....	168
10.5	Einschätzung der Herstellerkompetenz.....	169
10.6	Empfehlungen hinsichtlich einer Verwendung bei Personaldokumenten.....	169
	<b>Literatur</b> .....	<b>172</b>

---

## Abkürzungsverzeichnis

BDr	Bundesdruckerei
BGS	Bundesgrenzschutz
BKA	Bundeskriminalamt
BSI	Bundesamt für Sicherheit in der Informationstechnik
CBEFF	Common Biometric Exchange Format Framework
COG	Cognitec
DER	Dermalog
EBS	Einzelbenutzerstatistik
FAR	False Accept Rate (Falschakzeptanzrate)
FE	Fingerabdruckerkennung
FMR	False Match Rate
FNMR	False Non-Match Rate
FRR	False Reject Rate (Falschrückweisungsrate)
FTA	Failure-To-Acquire Rate
FTE	Failure-To-Enrol Rate
GE	Gesichtserkennung
ICAO	International Civil Aviation Organization
IE	Iriserkennung
MRZ	Machine Readable Zone
NIST	National Institute of Standards and Technology
ODBC	Open Database Connectivity
RefID1	Bild des Gesichts als Frontalaufnahme
RefID2	System-Template auf Basis Live-Enrolment am GE-System von Cognitec

---

RefID3	Bild eines Fingerabdrucks
RefID4	System-Template auf Basis Live-Enrolment am FE-System von BDr/NEC
RefID5	Bild eines Fingerabdrucks
RefID6	System-Template auf Basis Live-Enrolment am FE-System von Dermalog
RefID7	Bild einer Iris
RefID8	System-Template auf Basis Live-Enrolment am IE-System
SDI	SD Industries
UCL	University College London
VNC	Virtual Network Computing
VPN	Virtual Private Network
WSQ	Wavelet Scalar Quantization

## 1 Vorwort

Im Zuge der Maßnahmen zur Bekämpfung des internationalen Terrorismus wird eine Verbesserung der Identitätsprüfung in den verschiedenen Stufen der Einreise- und Aufenthaltskontrolle durch den Einsatz biometrischer Verfahren angestrebt. Als wesentliche Grundlage für eine Umsetzung entsprechender Überlegungen hat der Bundestag das am 09.01.2002 in Kraft getretene Terrorismusbekämpfungsgesetz beschlossen, das die Anpassung zahlreicher Sicherheitsgesetze an die neue Bedrohungslage beinhaltet. Zu den geänderten Bestimmungen zählen unter anderem das Passgesetz, das Gesetz über Personalausweise, das Ausländergesetz sowie das Asylverfahrensgesetz. Dabei werden insbesondere verschiedene Aspekte der Personenidentifizierung neu geregelt. So dürfen nun Pässe und Personalausweise neben dem Lichtbild und der Unterschrift weitere biometrische Merkmale in digitaler Form beinhalten.

Weitere Motivation für die Beschäftigung mit diesem Thema stellen die internationalen Aktivitäten und Rahmenbedingungen in diesem Umfeld dar. So hat der Kongress der USA ein Gesetzespaket zur Terrorismusbekämpfung beschlossen, das unter anderem wesentliche Änderungen des visafreien Einreiseprogramms („Visa Waiver Program“) beinhaltet. Hierbei wird von den teilnehmenden Staaten und damit auch Deutschland verlangt, dass bis Oktober 2006 biometrische Merkmale in deren Reisedokumente eingeführt werden. Entsprechend der Zielsetzung des US-amerikanischen „Enhanced Border Security and Visa Entry Reform Act“ vom 14.05.2002 auf Grundlage des „US Patriot Act“ für die Einführung von Biometrie in Reisedokumenten der Visa-Waiver-Staaten befasst sich auch die ICAO (International Civil Aviation Organization) mit Empfehlungen zur Erweiterung von Reisedokumenten um biometrische Merkmale. Die ICAO gibt grundsätzlich den Einsatz von Gesichtserkennung als das biometrische Merkmal für globale Interoperabilität vor, lässt jedoch optional weitere Merkmale wie Fingerabdrücke oder Iriden zu.

Auf europäischer Ebene wurde am 13.12.2004 durch den Rat der Europäischen Union die Verordnung über Normen für Sicherheitsmerkmale und biometrische Daten in von den Mitgliedsstaaten ausgestellten Pässen und Reisedokumenten erlassen. Danach sind Pässe und Reisedokumente mit einem Speichermedium zu versehen, das 18 Monate nach Verabschiedung der technischen Spezifikationen (d.h. bis Mitte 2006) ein Gesichtsbild enthält, und nach weiteren 18 Monaten sind Fingerabdruckbilder in interoperablen Formaten hinzuzufügen.

Am 01.06.2005 hat Bundesinnenminister Schily die Einführung neuer biometrieunterstützter Reisepässe - kurz: ePässe - in Deutschland vorgestellt. Damit führt Deutschland als einer der ersten EU-Staaten den neuen "EU-Reisepass" ein. Dieser neue ePass wird ab dem 01.11.2005 ein digitales Lichtbild enthalten, ab März 2007 zusätzlich zwei Fingerabdruckbilder.

Vor diesem Hintergrund war es Ziel der Studie BioP II, die Leistungsfähigkeit von derzeit auf dem Markt verfügbaren biometrischen Verifikationssystemen der Gesichts-, Fingerabdruck- und Iriserkennung für eine Verwendung im Zusammenhang mit Personaldokumenten zu untersuchen.

Dieser Bericht stellt die Testkonzeption und die wichtigsten Ergebnisse von BioP II vor. Die Studie wurde unter der Gesamtprojektleitung des BSI gemeinsam mit dem BKA in enger Kooperation mit den beiden Unternehmen Fraport AG und Deutsche Lufthansa AG bis Herbst 2004 durchgeführt und durch die Firma secunet Security Networks AG als Auftragnehmer realisiert.

Bonn, Wiesbaden, Essen, im August 2005



## 2 Management Summary

Der vorgesehene Einsatz biometrischer Verfahren zur Identitätsprüfung und die damit verbundene Aufnahme biometrischer Merkmale in Personaldokumente motiviert eine detaillierte Untersuchung entsprechender Technologien. Grundlage hierfür sind unter anderem die zahlreichen Gesetzesänderungen, die im Rahmen des Terrorismusbekämpfungsgesetzes vom 09.01.2002 im Bereich der Personaldokumente erfolgt sind und neben dem bisher traditionell eingesetzten Lichtbild und der Unterschrift eine Verwendung weiterer biometrischer Merkmale ermöglichen. Daneben verlangt das Visa-Waiver-Programm der USA, dass für den visafreien Reiseverkehr in die Vereinigten Staaten die Einreisenden über maschinell lesbare Reisedokumente mit biometrischen Merkmalen verfügen müssen. Des Weiteren ist im Dezember 2004 seitens des Rates der Europäischen Union die Aufnahme der biometrischen Merkmale von Gesicht und Finger in europäische Reisedokumente beschlossen worden.

In der Vorläuferstudie BioP I [BioPI], die 2003 abgeschlossen wurde, erfolgte die Untersuchung biometrischer Gesichtserkennungsverfahren, um daraus Erkenntnisse für eine Verwendung von Gesichtserkennung im Zusammenhang mit Personaldokumenten zu gewinnen. BioP I beinhaltete zum einen eine vergleichende Untersuchung zweier Gesichtserkennungssysteme und zum anderen den Vergleich mehrerer Gesichtserkennungsalgorithmen in unterschiedlicher Ausprägung. Darüber hinaus wurden auf Grundlage der vielfältigen in Betracht kommenden Personaldokumente mit Lichtbildern mehrere Varianten der Bereitstellung des Referenzbildes untersucht, um Aussagen darüber zu treffen, ob und mit welchen Ergebnissen die getesteten Systeme Lichtbilder unterschiedlichster Art und Qualität auf Personaldokumenten in einem Verifikationsprozess verarbeiten können.

Während somit in BioP I ausschließlich die Gesichtserkennung untersucht wurde, hatte BioP II im unmittelbaren Anschluss daran eine vergleichende Betrachtung der Verfahren Gesichtserkennung, Fingerabdruck- und Iriserkennung im praktischen Feldtest zum Ziel.

Die Untersuchung biometrischer Verifikationsverfahren im Rahmen der Studie BioP II diente dazu, die Leistungsfähigkeit von zum gegenwärtigen Zeitpunkt auf dem Markt verfügbaren Systemen zur Gesichtserkennung, Fingerabdruck- und Iriserkennung bezüglich verschiedener Aspekte zu analysieren und aus diesen Erkenntnissen Aussagen für eine erfolgreiche Verwendung im Zusammenhang mit Personaldokumenten zu entwickeln. Die Ergebnisse von BioP II sollten zudem Hilfestellung für die Art und Weise der Einführung der neuen ePässe geben.

Das Projekt BioP II beinhaltete eine aufgrund wissenschaftlicher Kriterien angelegte Praxiserprobung der drei genannten biometrischen Verfahren in einem vergleichenden Systemtest. Insbesondere im Hinblick auf die internationale Interoperabilität wurde analysiert, inwieweit der Einsatz ICAO- und ISO-konformer Bilder der biometrischen Merkmale als Vergleichsmuster (Referenzbasis) für die biometrische Verifikation praktikabel ist. Begleitend zum praktischen Systemtest wurden Benutzbarkeits- und Akzeptanzuntersuchungen durchgeführt.

Bei der Bewertung und Einordnung der hier erzielten Ergebnisse kommt es, wie bei anderen ähnlichen Untersuchungen, auf die stets begleitend zu berücksichtigenden Rahmenbedingungen sowie des gewählten Szenarios an. Das Herausgreifen einzelner Ergebnisse ohne Betrachtung des Gesamtzusammenhangs hat daher wenig Aussagekraft und kann zu Fehlinterpretationen führen. Hauptgegenstand der Untersuchungen in BioP II war zwar der oben bereits beschriebene Zusammenhang mit Personaldokumenten mit dem Ziel, für diese Anwendung verwertbare Aussagen für die drei biometrischen Verfahren zu erlangen. In einem Test kann dies jedoch nur auf Grundlage konkreter Produkte und Systeme erfolgen. Daher sind die hier erzielten Erkenntnisse zu den drei Verfahren nicht ohne weiteres verallgemeinerungsfähig und müssen stets anwendungsspezifisch betrachtet werden. Dies gilt vor allem im Hinblick auf die teilweise bedeutsamen Unterschiede der Produkte und Systeme, die lediglich grundsätzlich dasselbe biometrische Verfahren verwenden.

## ***Projektziele***

Inhalt des Projekts BioP II war die Untersuchung der in diesem Zusammenhang auftretenden Fragestellungen zur Machbarkeit und technischen Realisierung:

- Welche Erkennungsleistung bieten die Verfahren Gesichts-, Fingerabdruck- und Iriserkennung?
- Können die Verfahren mit Bilddateien gemäß ICAO als Referenz statt mit einer proprietären herstellerspezifischen Kodierung des Merkmals (Template) zufrieden stellende Ergebnisse erzielen?
- Welche Praxistauglichkeit bieten die drei Verfahren für eine große Testgruppe? Wie ist es um Benutzbarkeit und Akzeptanz bestellt?
- Welches der getesteten Verfahren eignet sich unter welchen Bedingungen am ehesten für die geplante Verwendung in Personaldokumenten?<sup>1</sup>

Bei der Bearbeitung dieser Fragestellungen wurden die internationalen Rahmenbedingungen – insbesondere die Richtlinien der ICAO und der ISO für die Aufnahme und Bereitstellung biometrischer Daten in Personaldokumenten – in der jeweils aktuellen Fassung berücksichtigt.

Für die genannten Zielvorgaben wurden im Rahmen von BioP II ein Gesichtserkennungssystem der Cognitec Systems GmbH, zwei Fingerabdruckerkennungssysteme der Hersteller Bundesdruckerei GmbH (in Zusammenarbeit mit NEC Deutschland) und Dermalog Identification Systems GmbH und ein Iriserkennungssystem der SD Industries GmbH (jetzt takeID GmbH) im Verifikationsmodus getestet. Die Entscheidung für das Gesichtserkennungssystem erfolgte anhand der Untersuchungen in BioP I. Die Auswahl der Fingerabdruck- und Iriserkennungssysteme wurde auf Basis eines im Vorfeld durchgeführten Auswahltests getroffen. Innerhalb dieser Voruntersuchung wurde eine Marktsichtung durchgeführt und mehrere Systeme entsprechend wesentlicher Anforderungen von BioP II analysiert. Insgesamt wurden Fingerabdrucksysteme von acht Anbietern und Iriserkennungssysteme von drei Anbietern untersucht.

## ***Vorgehen***

Bei der biometrischen Verifikation wird eine aktuelle Aufnahme des Merkmals (Gesicht, Fingerabdruck bzw. Iris) mit einer im Vorfeld gespeicherten Referenzaufnahme der entsprechenden Person, der so genannten Referenzbasis, verglichen. Die Bereitstellung der Referenzbasis erfolgte in digitaler Form. Hierfür kam sowohl eine Bilddatei des jeweiligen Merkmals als auch eine spezielle, herstellerspezifische Kodierung des Merkmals, das Template, zur Anwendung. Die Erstellung der Bilddateien und die Bereitstellung in einem interoperablen Format erfolgten gemäß den aktuellen Richtlinien von ICAO und ISO.

Um möglichst viele Betätigungen von Testteilnehmern für die Auswertung zu gewährleisten, fanden die Untersuchungen mittels eines umfangreichen Feldtests an vier verschiedenen Standorten am Flughafen Frankfurt/Main statt. Dabei wurden die Systeme in einem Zeitraum von vier Monaten von etwa 2000 Mitarbeitern der Fraport AG, der Deutsche Lufthansa AG sowie des Bundesgrenzschutzes (BGS) benutzt. Durch die Einschränkung der Testgruppe auf Berufstätige war keine Repräsentativität für die Gesamtbevölkerung der Bundesrepublik Deutschland gegeben. Alle Testteilnehmer waren

---

<sup>1</sup> Diese Fragestellung wurde teilweise von den nationalen und europäischen Entwicklungen während der bereits laufenden Untersuchungen überholt, da zwischenzeitlich die Entscheidungen, wie im Vorwort benannt, getroffen wurden.

angehalten, über die Dauer des Feldtests mindestens zweimal täglich einen der Teststandorte aufzusuchen und dort an zwei unterschiedlichen Systemen je einen Verifikationsversuch durchzuführen.

Der Prozess war für jeden Teilnehmer wie folgt ausgestaltet:

Jedem Teilnehmer wurde beim Enrolment eine kontaktlose Smartcard (RF-Token) mit seiner Identifikationsnummer (UserID) ausgehändigt. Unter Verwendung dieser Smartcard teilten die Teilnehmer den Systemen ihre Identität mit. Anhand der auf der Smartcard gespeicherten UserID wurden die zugehörigen Referenzdaten (sowohl ICAO-Bilder als auch System-Templates) für biometrische Verifikationen aus einer Datenbank bereitgestellt. Die Verwendung einer Datenbank zur Bereitstellung der Daten diente hierbei lediglich testinternen Zwecken. Dies erfolgte unter Berücksichtigung der im Personalausweis- und Passgesetz vorgesehenen Speicherung der Referenzdaten im Dokument wie z.B. dem ePass.

Der konkrete Ablauf für die Bedienung eines Systems durch einen Testteilnehmer und die Durchführung der biometrischen Verifikation im späteren Testbetrieb stellte sich nach dem Enrolment folgendermaßen dar:

Die Person präsentierte dem biometrischen System mittels einer kontaktlosen Karte (RF-Token) eine Identifikationsnummer (UserID). Anhand dieser UserID wurden die biometrischen Referenzdaten aus einer Datenbank geladen. Anschließend erfolgte die Merkmalsaufnahme. Dabei wurden durch eine Sensoreinheit Bilder des jeweiligen biometrischen Merkmals aufgenommen und einem Vergleich mit einer gespeicherten Referenzbasis unterzogen. Je nachdem, ob der Versuch erfolgreich war oder nicht, wurde dies der Person über ein grünes bzw. rotes Lichtsignal mitgeteilt. Im Hintergrund und für den Teilnehmer nicht wahrnehmbar erfolgten weitere Vergleiche des aufgenommenen Bildes mit der zweiten Referenzbasis. Alle erzielten Ergebnisse wurden für die spätere Auswertung in einer Datenbank protokolliert. Die biometrische Verifikation war nicht mit einer konkreten Anwendung (wie zum Beispiel Zutrittskontrolle) verbunden, sondern stellte ein reines Testszenario dar.

Neben dem beschriebenen Feldtest wurden eine Reihe von weiterführenden Untersuchungen im Labor der Firma secunet Security Networks AG in Essen durchgeführt. Wesentliche Bestandteile waren die Prüfung der Überwindungssicherheit der beteiligten Systeme sowie die Durchführung verschiedener datenbankbasierter Tests zur Untersuchung der Leistungsfähigkeit der biometrischen Verfahren.

Zur Darstellung und zum Vergleich der gewonnenen Ergebnisse wurden diese klassifiziert, anhand eines Notensystems bewertet und gemäß ihrer Bedeutung für die Anwendung in Personaldokumenten gewichtet in Bewertungsschemata abgebildet. Die Zusammenfassung der gewichteten Teilnoten resultierte in Gesamtnoten, die eine vergleichende Bewertung ermöglichen.

Die wesentlichen Ergebnisse sind in den folgenden Abschnitten dargestellt. Insbesondere werden dabei folgende Punkte betrachtet:

- Erkennungsleistung
- Benutzbarkeit und Akzeptanz
- Überwindungssicherheit<sup>2</sup>
- Speicherung biometrischer Daten auf RF-Token

---

<sup>2</sup> Auf eine detaillierte Darstellung wird in diesem Bericht verzichtet.

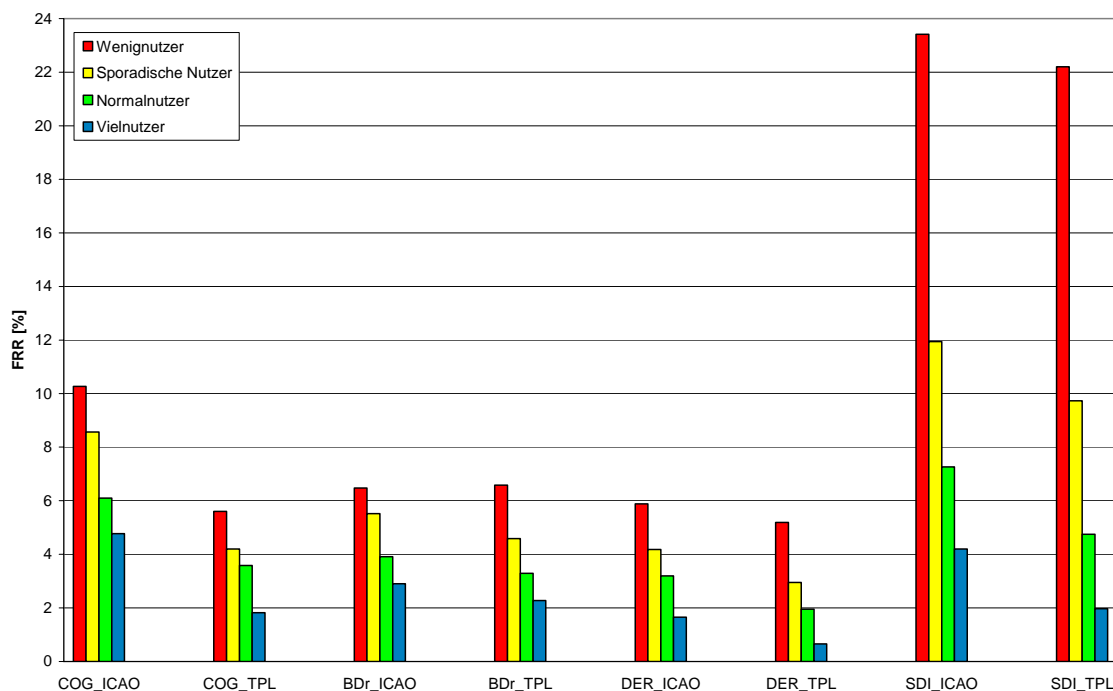
## ***Erkennungsleistung***

Die Bewertung der Erkennungsleistung muss stets anwendungsspezifisch erfolgen. Wie stark die klassischen Fehlerraten Falschakzeptanz (FAR) und Falschrückweisung (FRR) neben weiteren Bewertungskriterien gewichtet werden, hängt entscheidend von der konkreten Anwendung ab, für die ein biometrisches System eingesetzt werden soll. Während in einem sicherheitssensiblen Bereich die Optimierung des Systems auf eine möglichst geringe FAR zur Erreichung einer hohen Sicherheit angezeigt ist, kann für ein anderes Szenario etwa mit zusätzlicher manueller Kontrolle das System eher auf den Benutzungskomfort und damit auf eine möglichst geringe FRR optimiert werden. Das bei BioP II zu berücksichtigende Anwendungsszenario "Personaldokumente" hat eine entsprechende Gewichtung der definierten Bewertungskriterien zur Folge, auf die in Kapitel 9 (Bewertungsschema) näher eingegangen wird.

Zur Bewertung der Erkennungsleistung wurden die klassischen Fehlerraten FRR (Falschrückweisungsrate) und FAR (Falschakzeptanzrate) ermittelt. Hierfür wurden ausgehend von verschiedenen Sicherheitsniveaus, welche durch die FAR repräsentiert werden, die zugehörigen FRRs berechnet. Zudem wurde auch die FTE (Failure-to-Enrol Rate) bewertet.

Die FTE, also der Anteil von Personen, die an einem biometrischen System nicht erfasst werden können, war für alle beteiligten Systeme relativ gering. Bei der Gesichtserkennung konnten alle Personen enrolt werden, bei der Fingerabdruck- und Iriserkennung ist jeweils bei weniger als 1% der teilnehmenden Personen das Enrolment gescheitert. Dies ist eine im Vergleich zu den Erwartungen der Hersteller und Erfahrungen aus anderen Tests sehr niedrige Quote.

Die Testteilnehmer bedienten die biometrischen Systeme mit sehr unterschiedlichen Häufigkeiten. Daher erfolgte eine Einteilung der Teilnehmer in Nutzerklassen je System gemäß der Häufigkeit der Benutzung über den gesamten Testzeitraum: „Wenignutzer“ (weniger als zehn Betätigungen), „Sporadische Nutzer“ (zehn bis 30 Betätigungen), „Normalnutzer“ (30 bis 120 Betätigungen) und „Vielnutzer“ (mehr als 120 Betätigungen). Die Erkennungsleistung war für alle getesteten Verfahren abhängig von der Nutzerklasse. Personen, die nur wenige Betätigungen durchführten, verursachten höhere Rückweisungsrate als geübte Nutzer (vgl. Abbildung 1: Nutzerklassen-spezifische Erkennungsleistung). Dieser Effekt war am ausgeprägtesten bei dem untersuchten Iriserkennungssystem. Dieser Effekt wird sich durch eine verbesserte Benutzerführung am Gerät erheblich reduzieren lassen. Auch im Hinblick auf die unterschiedlichen Nutzerklassen muss also ein biometrisches System anwendungsspezifisch beurteilt werden. Während etwa bei einem System, das typischerweise sehr häufig und von Vielnutzern bedient wird, durch den schnell erzielbaren Trainingseffekt vermutlich niedrige FRRs erreichbar sind, ist für eine Anwendung mit seltenen Nutzern, aufgrund deren Ungeübtheit mit höheren Rückweisungsrate zu rechnen. Auch für diese Nutzergruppe ist allerdings bei zunehmender Nutzung biometrischer Systeme allgemein mit verbesserten Rückweisungsrate zu rechnen.



**Abbildung 1: Nutzerklassen-spezifische Erkennungsleistung (FAR=0,1%)<sup>3</sup>**

Bei der Betrachtung der Entwicklung der Fehlerraten über die Laufzeit des Feldtests wurde deutlich, dass der Trainingseffekt für die verschiedenen Systeme durchaus unterschiedlich ist. Während sich bei dem Gesichtserkennungssystem nach einer bestimmten Zeit die FRR stabilisierte, war noch zum Ende des Feldtests bei beiden Fingerabdruckerkennungssystemen weiterhin ein schwacher und beim Iriserkennungssystem ein deutlicher Rückgang der FRR zu verzeichnen. Insbesondere für das Iriserkennungssystem ist daher ein erhebliches Verbesserungspotenzial durch intensives Training der Nutzer zu erwarten. Die Schwierigkeiten beim Iriserkennungssystem sind hier im Wesentlichen auf die verwendete Kameraeinheit zurückzuführen. Mit einer anderen Sensoreinheit, die auch für ungeübte Nutzer eine einfache Merkmalspräsentation ermöglicht und über eine geeignete Benutzerführung verfügt, ist eine deutliche Verbesserung der Erkennungsleistung zu erwarten. Darüber hinaus wird der zunehmende Einsatz biometrischer Systeme im Alltag zu einer Gewöhnung und Übung der Nutzer führen, die diesen Effekt der verbesserten Erkennungsleistung noch verstärken werden.

Die durchgeführten Untersuchungen haben anhand der hier eingesetzten Systeme gezeigt, dass bei einer Erhöhung des Sicherheitsniveaus ( $FAR < 0,1\%$ ) die Erkennungsrate der Gesichtserkennung stärker abnimmt als bei der Fingerabdruck- und Iriserkennung.

<sup>3</sup> COG = Gesichtserkennungssystem Cognitec

BDr = Fingerabdruckerkennungssystem Bundesdruckerei/NEC

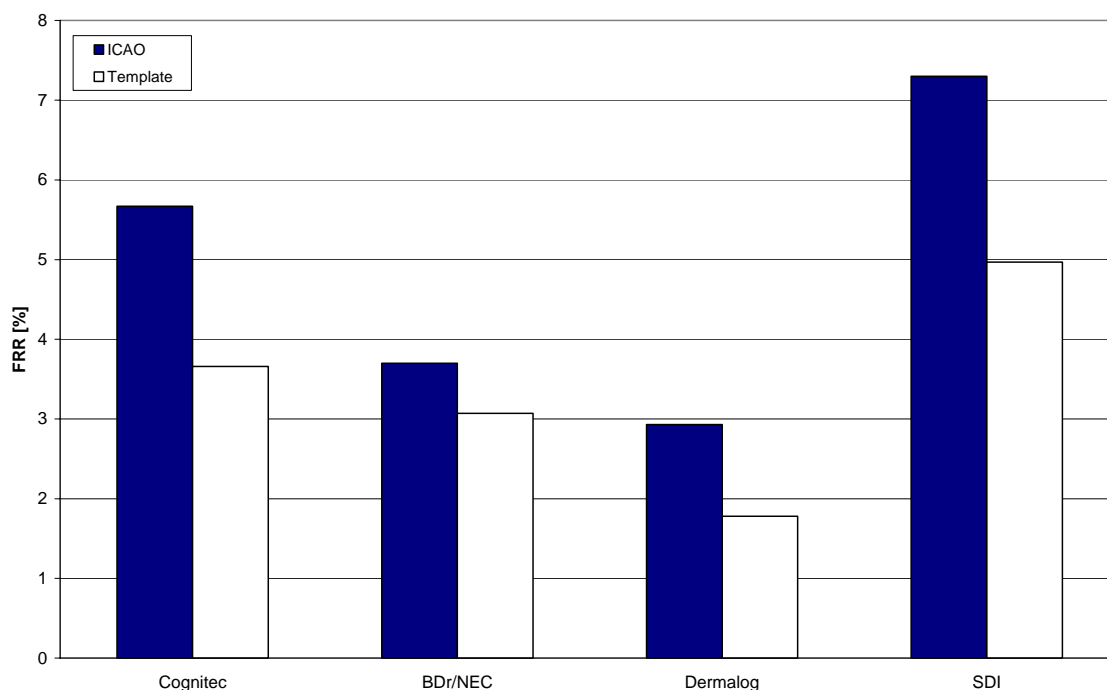
DER = Fingerabdruckerkennungssystem Dermalog

SDI = Iriserkennungssystem SD Industries

ICAO = Referenzbasis ICAO-Bild

TPL = Referenzbasis Systemtemplate

Die Erkennungsleistung bei der Verwendung von ICAO-Referenzbildern war niedriger als bei Verwendung von Referenz-Templates, die beim System-Enrolment aufgenommen wurden. Dies kann damit begründet werden, dass die zurzeit verfügbaren biometrischen Systeme in der Regel auf die herstellereigenen proprietären Templates trainiert sind. Die Erkennungsleistung mit ICAO-Referenzbildern bewegte sich jedoch immer noch in einer akzeptablen Größenordnung (vergleiche Abbildung 2). Ein Einsatz biometrischer Verfahren mit ICAO-Bildern als Referenz ist daher bereits heute mit zufriedenstellenden Ergebnissen möglich. Weitere Verbesserungen sind durch die Entwicklungen in der Standardisierung biometrischer Verfahren zu erwarten. Bereits heute vorhandene Optimierungspotentiale der Algorithmen bei der Verarbeitung entsprechender Bilddaten können so weiter ausgebaut werden.



**Abbildung 2: Erkennungsleistung in Abhängigkeit der Referenzbasis (FAR=0,1%)**

Bei der Betrachtung der in BioP II ermittelten Falschrückweisungsraten muss berücksichtigt werden, dass mit der biometrischen Verifikation keine echte Anwendung (wie beispielsweise eine Zutrittskontrolle) verknüpft war, es sich also um einen reinen Testbetrieb handelte. In realen Einsatzumgebungen ist aufgrund der zu erwartenden erhöhten Kooperationsbereitschaft mit geringeren Rückweisungsraten zu rechnen, da hier in aller Regel davon auszugehen ist, dass die Nutzer am Erfolg der Verifikation ein unmittelbares Interesse haben. Des Weiteren werden biometrische Systeme in der Praxis in aller Regel in der Weise betrieben, dass die Möglichkeit mehrerer direkt aufeinander folgender Versuche für die Identitätsprüfung besteht. Auf diese Einstellung wurde bei BioP II aus Gründen der Testmethodik verzichtet und nur die jeweils erste Betätigung eines Versuchs ausgewertet. Durch die Kumulation mehrerer Versuche ist eine Verbesserung der FRR zu erwarten.

### ***Benutzbarkeit und Akzeptanz***

Die Testteilnehmer von BioP II standen dem Einsatz von Biometrie grundsätzlich sehr positiv gegenüber. Dies drückte sich in einer Zustimmung zum Einsatz biometrischer Verfahren für verschiedenste Anwendungszwecke aus. Gleichzeitig wurde großer Informationsbedarf u.a.

hinsichtlich der Aspekte der Erfassung (Gesundheit, Hygiene) und der Verarbeitung bzw. Speicherung (Datenschutz) biometrischer Daten festgestellt. Es handelte sich bei BioP II um freiwillige Testteilnehmer sowie solche Personen, die ihren eigenen Angaben nach gegenüber biometrischen Verfahren grundsätzlich kritisch eingestellt waren. Zudem wurden auch solche Personen befragt, die sich gegen eine Teilnahme am Test entschieden hatten.

Es konnte festgestellt werden, dass bei allen eingesetzten Systemen das Design und das Benutzerfeedback verbesserungswürdig sind. Dabei ist allerdings auch zu berücksichtigen, dass den Herstellern nur ein stark begrenzter Zeitraum für die Konstruktion der Systeme unter Berücksichtigung der Vorgaben des Auftraggebers zur Verfügung stand. Die Gestaltung der Systeme sollte unter anderem so erfolgen, dass der Benutzer dabei unterstützt wird, sein biometrisches Merkmal gegenüber dem System und Sensor möglichst optimal, d.h. für eine möglichst geringe Falschrückweisungsrate, zu präsentieren. Während bei Vielnutzern eine Benutzerführung kaum notwendig ist (hinsichtlich der dann unter Umständen längeren Bedienzeit ggf. sogar kontraproduktiv für die Akzeptanz), muss bei Wenignutzern auf diesen Aspekt sehr viel Wert gelegt werden. Bei dieser Gruppe ist eine ggf. längere Bedienzeit weniger relevant, wenn die Bedienung trotz seltener Benutzung problemlos und einfach erfolgen kann.

Abgesehen von Problemen, die aufgrund des Systemdesigns (zum Beispiel unklares Feedback, mangelnde Performance) auftraten, waren bei dem Gesichtserkennungssystem keine grundsätzlichen Schwierigkeiten bei der Präsentation des Merkmals zu beobachten. Bei den Fingerabdruckerkennungssystemen traten teilweise Probleme bei der Merkmalspräsentation hinsichtlich Auflagedruck und Positionierung auf. Da das in BioP II eingesetzte Iriserkennungssystem eine sehr genaue Positionierung des Auges vor der Erfassungseinheit erforderte, führte dies insbesondere bei ungeübten Personen zu Schwierigkeiten bei der Erkennung.

### ***Überwindungssicherheit***

Während der Fokus von BioP II auf der Erkennungsleistung lag, wurde sekundär auch die Überwindungssicherheit der untersuchten biometrischen Systeme beurteilt. Bei den hier erfolgten Überwindungsversuchen handelt es sich ganz überwiegend um Labortests, die nur vorläufige Aussagen zulassen. Auf eine detaillierte Darstellung wird in diesem Bericht daher verzichtet.

### ***Speicherung biometrischer Daten auf RF-Token***

Für die Integration biometrischer Daten in Personaldokumenten müssen diese über geeignete Speichermöglichkeiten verfügen. Seitens der ICAO wird die Verwendung kontaktloser Chips (RF-Token) vorgegeben. Da über die Projektlaufzeit mehrere Anpassungen der relevanten Standards der ICAO erfolgten, wurden die vorgesehenen Untersuchungen in eine separate Projektreihe „ILSE“ ausgelagert, die aktiv die Standardisierungsmaßnahmen der ICAO begleitet. Die Projektreihe ILSE wird analog zu BioP II in Zusammenarbeit von BSI, BKA und secunet durchgeführt.

Es existieren mittlerweile kontaktlose Smartcards und entsprechende Betriebssysteme, auf deren Basis die Speicherung biometrischer Daten erfolgen kann. Dabei sind technische Einschränkungen beispielsweise hinsichtlich des verfügbaren Speicherplatzes und erreichbarer Übertragungszeiten zu berücksichtigen.

Zum Schutz der biometrischen Daten sind Sicherheitsmechanismen auf verschiedenen Stufen umsetzbar. Diese unterscheiden sich im Wesentlichen in der Stärke der Authentisierung und orientieren sich am Schutzbedarf der gespeicherten Daten.

## ***Bewertung der getesteten Systeme***

Die **Gesichtserkennung** ist mit dem hier verwendeten System einfach benutzbar. Die Präsentation des Merkmals stellt für die Nutzer in der Regel kein Problem dar, so dass kaum Training erforderlich ist. Insbesondere für Anwendungen mit Wenignutzern ist dies ein Vorteil.

Im Hinblick auf den Einsatz in Personaldokumenten ist von Vorteil, dass bereits in der herkömmlichen Form Lichtbilder verwendet werden. Für eine digitale Verwendung muss zusätzlich sichergestellt werden, dass die verwendeten Bilder eine für die Gesichtserkennung geeignete Qualität aufweisen.

BioP II hat die Erfahrungen aus BioP I bestätigt und damit die Abhängigkeit der Gesichtserkennung von den Lichtbedingungen in der Umgebung. Bereits kleine Änderungen der Beleuchtungsverhältnisse haben bei dem hier eingesetzten System zu starken Änderungen der Erkennungsleistung geführt. Mit alternativen Erfassungseinheiten können vermutlich andere Ergebnisse erzielt werden. Insbesondere die Verwendung von ICAO-Bildern als Referenz erfordert eine homogene Ausleuchtung des Gesichts bei der Aufnahme des Live-Bildes. Bei sehr unterschiedlichen Bedingungen an potenziellen Einsatzorten ist daher mit verschiedenen Resultaten in der Erkennungsleistung zu rechnen. Auch beim Einsatz von Templates muss für vergleichbare Lichtbedingungen bei Enrolment und Verifikation gesorgt werden.

Schließlich kann Gesichtserkennung grundsätzlich von allen Ausweisinhabern genutzt werden, es können alle Gesichtsbilder erfasst werden. Ausnahmen bilden Personen, für die sich etwa aus religiösen bzw. kulturellen Gründen eine Aufnahme des Gesichts verbietet, oder aus anderen Gründen wesentliche Teile des Gesichts verdeckt sind.

Die **Fingerabdruckerkennung** liefert mit den beiden hier untersuchten Systemen im Vergleich zu den beiden weiteren Verfahren die beste Erkennungsleistung. Beide Fingerabdruckererkennungssysteme können bei den hier betrachteten Sicherheitsniveaus die guten Erkennungsraten der beiden anderen Verfahren noch überbieten. Die Ergebnisse zeigen damit, dass die Fingerabdruckerkennung auch für den Einsatz bei sehr hohen Sicherheitsniveaus geeignet ist.

Bei beiden Fingerabdruckererkennungssystemen kamen im Feldtest optische Sensoren zum Einsatz. Trotz der hohen Frequentierung (bis zu 18.000 Betätigungen pro Sensor) wiesen die Sensoren eine sehr hohe Stabilität auf. Es konnte bei mehrfach täglicher Reinigung weder eine nennenswerte Verschlechterung der Bildqualität beobachtet werden, noch gab es Gebrauchs- oder Verschleißspuren.

Aufgrund z.B. schwacher Merkmalsausprägung oder körperlichen Einschränkungen, die eine geeignete Präsentation des Merkmals verhindert, kann ein gewisser Anteil potenzieller Nutzer die Fingerabdruckerkennung nicht benutzen. Der Anteil dieser nicht-enrolbaren Personen könnte allerdings reduziert werden, indem beispielsweise der Daumen verwendet wird, der in der Regel über eine höhere Anzahl von auswertbaren Merkmalen verfügt als andere Finger.

Die Erkennungsleistung des hier untersuchten **Iriserkennungssystems** würde den Einsatz für hohe Sicherheitsniveaus ermöglichen. Es ist festzuhalten, dass die Falschakzeptanzrate (FAR) für die fest implementierte Toleranzschwelle höher ist als von den Herstellern regelmäßig angegeben. Geeigneter erscheint daher, den fest implementierten Schwellwert nicht als Dogma anzusehen, sondern auch hier Konfigurationsmöglichkeiten hinsichtlich verschiedener Sicherheitsniveaus zu schaffen.




Festzustellen waren bei dem hier getesteten Iriserkennungssystem Bedienungsschwierigkeiten für ungeübte Benutzer. Da das eingesetzte Kamerasystem eine genaue Positionierung des Auges erfordert, führte dies bei Personen, die das System nur selten benutzten, gehäuft zu Problemen, die sich für in einer FRR von über 20% in der Gruppe der Wenignutzer widerspiegeln. Dieser Effekt war jedoch

auch bei einzelnen Personen festzustellen, die die Iriserkennungssysteme regelmäßig benutzen. Mit einer anderen Sensoreinheit, die auch für ungeübte Nutzer eine einfache Merkmalspräsentation ermöglicht und über eine geeignete Benutzerführung verfügt, ist eine deutliche Verbesserung der Erkennungsleistung zu erwarten. Darüber hinaus wird der zunehmende Einsatz biometrischer Systeme im Alltag zu einer Gewöhnung und Übung der Nutzer führen, die zu einer weiter verbesserten Erkennungsleistung führen werden.

Auch bei der Iriserkennung ist grundsätzlich zu berücksichtigen, dass ein gewisser Anteil von Personen diese nicht nutzen können, weil diesen z.B. aufgrund starker Sehschwäche oder anderer körperlicher Einschränkungen eine geeignete Präsentation der Iris nicht möglich ist.

### ***Vergleichende Bewertung***

Aufgrund der im Rahmen der Untersuchungen erzielten Ergebnisse ist eine Bewertung der getesteten Systeme mittels verschiedener Kriterien möglich. Diese Bewertung ist, da anhand konkreter biometrischer Systeme in einem bestimmten Anwendungsszenario getroffen, nicht ohne weiteres verallgemeinerungsfähig. Die hier erzielten Ergebnisse zeigen jedoch Tendenzen der aktuellen Leistungsfähigkeit der betroffenen Verfahren auf und können deshalb zur Beantwortung wesentlicher Fragen für den hier relevanten Einsatzzweck herangezogen werden. Ausgehend vom angestrebten Einsatzzweck sind die Bewertungskriterien unterschiedlich zu gewichten. Für das Szenario Identitätskontrolle in Zusammenhang mit Personaldokumenten ergibt sich anhand der hier untersuchten Systeme ein Vorsprung der Fingerabdruckerkennung vor den beiden anderen Technologien, von denen die Gesichtserkennung etwas besser abschneidet als die Iriserkennung.

	<p><b>1</b></p> <ul style="list-style-type: none"> <li>+ Beste Erkennungsleistung</li> <li>+ Hohes Sicherheitsniveau möglich</li> <li>+ Stabile Sensoren trotz kontaktbehaftetem Einsatz</li> </ul>
	<p><b>2</b></p> <ul style="list-style-type: none"> <li>+ Kaum Training erforderlich</li> <li>- Ausgesprochen homogene Ausleuchtung für ICAO-Bilder nötig</li> <li>- Bei hohem Sicherheitsniveau nicht akzeptable Rückweisungsrate</li> </ul>
	<p><b>3</b></p> <ul style="list-style-type: none"> <li>+ Hohes Sicherheitsniveau möglich</li> <li>- Bedienungsschwierigkeiten für Wenignutzer</li> <li>- FAR signifikant höher als angegeben ("Mythos feste Schwelle")</li> </ul>

**Abbildung 3: Zusammenfassender Vergleich der untersuchten Systeme**

Der Vorsprung der Fingerabdruckerkennungssysteme ergibt sich durch die insgesamt bessere Erkennungsleistung. Die Abwertung des Iriserkennungssystems resultiert aus den hohen Falschrückweisungsraten für ungeübte Nutzer.

### ***Empfehlungen hinsichtlich einer Verwendung bei Personaldokumenten***

Biometrische Verfahren können die Identitätsprüfung anhand von Personaldokumenten wirksam unterstützen. Hierbei sind die unterschiedlichen Anforderungen und Umweltbedingungen verschiedener Szenarien zu beachten (z.B. verschiedene Grenzübertritte, personenkontrolle durch die Polizei etc.). Neben technischen Maßnahmen erscheinen daneben auch organisatorische Maßnahmen geeignet, eine für die Ziele der Anwendung geeignete Erkennungsleistung in einer konkreten Anwendung unter Berücksichtigung sowohl der Falschakzeptanz- als auch der Falschzurückweisungsrate zu erreichen. Darüber hinaus sollte der Benutzer bei der Präsentation des Merkmals durch eine geeignete Benutzerführung unterstützt werden.

Für eine geeignete Bereitstellung der Referenzdaten müssen Personaldokumente um ein Speichermedium – einen kontaktlosen Chip gemäß den Vorgaben der ICAO – erweitert werden. Neben einer entsprechenden Speicherkapazität muss dieser auch über Sicherheitsmechanismen verfügen, die einen unberechtigten Zugriff auf die biometrischen Daten des Ausweisinhabers verhindern.

Eine wesentliche Voraussetzung für den erfolgreichen Einsatz biometrischer Verfahren ist die Kooperation und somit die Akzeptanz seitens der Benutzer. Aufgrund der hier gewonnenen Erkenntnisse besteht Bedarf an Aufklärung und Information sowie weiteren akzeptanzfördernden Maßnahmen.

Die Benutzerfreundlichkeit und Bedienbarkeit der hier untersuchten biometrischen Systeme ist nach den erzielten Ergebnissen verbesserungsfähig. Insbesondere müssen das Systemdesign und die Benutzerführung an der konkreten Anwendung und der Nutzergruppe ausgerichtet sein. Vor allem in Szenarien ohne zusätzliche manuelle Kontrolle sollte eine klare und einfache Benutzerführung die Bedienung des Systems so einfach und intuitiv wie möglich machen.

Beim Einsatz von Gesichtserkennung ist – wie bereits in BioP I aufgezeigt wurde – eine Anpassung der derzeitigen Aufnahmeleitlinien zur Lichtbilderstellung erforderlich. Hier stellen die aktuellen Richtlinien der ICAO für die Erstellung von Passbildern zum Einsatz für biometrische Anwendungen eine geeignete Grundlage dar. Daneben ist unter anderem darauf zu achten, dass wesentliche Gesichtspartien nicht verdeckt sind (z.B. durch Kopfbedeckung, Frisur oder ähnliches).

Die Gesichtserkennung stellt hohe Anforderungen an die Einsatzumgebung. Hinsichtlich der Lichtbedingungen muss eine hohe Homogenität sichergestellt werden.

Weitere Verbesserungen hinsichtlich Erkennungsleistung und Sicherheit können möglicherweise durch geeignete Kombination verschiedener biometrischer Verfahren oder Algorithmen erreicht werden. Die Möglichkeiten und Leistungsfähigkeit solcher multimodaler Ansätze werden zurzeit geprüft.

Ebenso ist zu beachten, dass der Einfluss von Alterungseffekten auf die Erkennungsleistung biometrischer Systeme bisher noch nicht ausreichend untersucht ist. Dies ist insbesondere vor dem Hintergrund des langen Gültigkeitszeitraums von Personaldokumenten von bis zu zehn Jahren zu berücksichtigen.

Schließlich sollte zudem darauf geachtet werden, dass die Systeme hohe Anforderungen bzgl. der IT-Sicherheit sowie der physischen Sicherheit erfüllen müssen, um einen geeigneten Manipulationsschutz und Schutz der persönlichen Daten zu gewährleisten. Darüber hinaus müssen die Systeme in eine sichere Gesamtumgebung integriert werden.

Abschließend erscheint vor dem Echtbetrieb in einer konkreten Anwendung eine gründliche Untersuchung der Funktionstüchtigkeit, der Erkennungsleistung und der Überwindungssicherheit sinnvoll und notwendig. Dies kann etwa durch eine Pilotierung im Echtbetrieb unter Berücksichtigung der praktisch relevanten Parameter für die konkrete Anwendung erfolgen, um ggf. noch notwendige

---

Anpassungen vor der tatsächlichen Inbetriebnahme in der Praxis ohne unverhältnismäßigen Aufwand vornehmen zu können.

BioP II hat das aktuelle Leistungspotenzial der drei biometrischen Verfahren Gesichts-, Fingerabdruck- und Iriserkennung anhand konkreter Systeme aufgezeigt sowie Verbesserungsmöglichkeiten hinsichtlich der Erkennungsleistung einerseits und der Benutzbarkeit andererseits identifiziert. Die hier erarbeiteten Empfehlungen können als erste Richtschnur für die Verwendung biometrischer Verfahren in Verbindung mit Personaldokumenten dienen.



## 3 Einleitung

### 3.1 Testziel

Mit dem am 09.01.2002 in Kraft getretenen Terrorismusbekämpfungsgesetz und der damit einhergehenden Anpassung des Pass- und Personalausweisgesetzes werden insbesondere verschiedene Aspekte der Personenidentifizierung neu geregelt.

Die Neuregelung umfasst unter anderem die Erweiterung der Personaldokumente um biometrische Merkmale zur Verbesserung des Prozesses der Identitätsfeststellung. Hierbei gilt, dass keine zentrale Speicherung dieser biometrischen Merkmale erfolgen darf. Daher bietet sich eine Verwendung von biometrischen Verfahren im Verifikationsmodus (1:1-Vergleich eines im Personaldokument gespeicherten Merkmals gegen das entsprechende Live-Merkmal der Person) an.

Weitere Motivation für die Beschäftigung mit diesem Thema stellen die internationalen Aktivitäten und Rahmenbedingungen in diesem Umfeld dar. So hat der Kongress der USA ebenfalls ein Gesetzespaket zur Terrorismusbekämpfung beschlossen, das unter anderem wesentliche Änderungen des visafreien Einreiseprogramms („Visa Waiver Program“) beinhaltet. Hierbei wird von den teilnehmenden Staaten – zu diesen gehört auch Deutschland – verlangt, dass biometrische Merkmale in deren Reisedokumente eingeführt werden oder dass zumindest ein Programm mit entsprechenden Plänen nachgewiesen wird. Als Folge dieser Vorschriften unternehmen derzeit zahlreiche Staaten erhebliche Anstrengungen zur Einführung biometrischer Merkmale in ihre Reisedokumente.

Nicht zuletzt durch die zahlreichen internationalen Aktivitäten wird deutlich, dass eine Abstimmung mit anderen Staaten erforderlich ist, um eine Vielzahl nicht kompatibler Lösungen zu vermeiden. Aus diesem Grund befasst sich die ICAO (International Civil Aviation Organization) intensiv mit Empfehlungen zur Erweiterung von Reisedokumenten um biometrische Merkmale. Die ICAO gibt grundsätzlich den Einsatz von Gesichtserkennung vor, lässt jedoch optional weitere Merkmale wie Fingerabdruck oder Iris zu.

Für die Integration biometrischer Merkmale in deutsche Personaldokumente müssen damit mehrere Aspekte berücksichtigt werden. Grundvoraussetzung für eine international interoperable Lösung ist die Beachtung der ICAO-Rahmenbedingungen, wobei zusätzlich eine Entscheidung für eine national bevorzugte Technologie getroffen werden kann. Dies beinhaltet einerseits die Festlegung eines biometrischen Merkmals sowie andererseits dessen geeignete Bereitstellung innerhalb des Dokuments. Außerdem muss eine Entscheidung für geeignete Anbieter dieser Technologie getroffen werden.

Damit die zuständigen Entscheidungsgremien geeignete Beschlüsse zur Aufnahme biometrischer Merkmale in Personaldokumente treffen können, müssen fundierte Bewertungskriterien auf Basis technischer und wissenschaftlicher Analysen bereitgestellt werden. Neben der generellen Eignung biometrischer Verfahren ist dabei auch die Marktsituation für die jeweilige Technologie zu untersuchen.

Innerhalb des Projekts BioP II wurden in einem umfangreichen Feldtest mit mehr als 2000 freiwilligen Testpersonen die biometrischen Verfahren Gesichtserkennung, Iriserkennung und Fingerabdruckerkennung in einer dem Zielszenario angelehnten Testumgebung vergleichend untersucht. Neben dem Feldtest wurden in weiterführenden Untersuchungen spezielle Fragestellungen wie zum Beispiel die Akzeptanz durch die Benutzer, die Benutzbarkeit der biometrischen Systeme sowie deren Überwindungssicherheit betrachtet.

Wesentliche Fragestellungen von BioP II sind:

- Welche Erkennungsleistung bieten die Verfahren Gesichtserkennung, Fingerabdruck- und Iriserkennung?
- Können die Verfahren mit Referenz-Bilddateien gemäß ICAO an Stelle von proprietären Referenz-Templates zufrieden stellende Ergebnisse erzielen?
- Welche Praxistauglichkeit bieten die drei Verfahren für eine große Testgruppe? Wie ist es um Benutzbarkeit und Akzeptanz bestellt?
- Welches der getesteten Verfahren eignet sich unter welchen Bedingungen am ehesten für die geplante Verwendung in Personaldokumenten?<sup>4</sup>

### 3.2 Systemauswahl

Zielsetzung des Projekts BioP II war eine vergleichende Untersuchung der Technologien Gesichtserkennung, Iris- und Fingerabdruckerkennung. Als Grundlage dazu wurde eine Festlegung auf ein zu testendes Gesichtserkennungs-, ein Iriserkennungs- und zwei Fingerabdruckerkennungssysteme getroffen. Die Entscheidung für zwei Fingerabdruckerkennungssysteme begründete sich in der großen Anzahl unterschiedlicher verfügbarer Sensoren und Algorithmen für diese Technologie.

Für die Gesichtserkennung wurde der Testsieger aus BioP I [BioPI], ein Komplettsystem der Firma Cognitec, eingesetzt. Die Auswahl der Systeme für Iris- und Fingerabdruckerkennung basiert auf einer durch das Institut für biometrische Identifikationssysteme (IBIS) durchgeführten Voruntersuchung für BioP II. Innerhalb dieser Voruntersuchung wurde eine Marktsichtung durchgeführt und einige Systeme entsprechend wesentlicher Anforderungen von BioP II analysiert. Insgesamt wurden Fingerabdruckerkennungssysteme von acht Anbietern sowie Iriserkennungssysteme von drei Anbietern untersucht.

Für die Iriserkennung wurde vom IBIS ein System von SD Industries empfohlen. Für die Fingerabdrucktechnologie wurden seitens IBIS Empfehlungen für drei Hersteller ausgesprochen: Bundesdruckerei/NEC, Dermalog und Intraproc.

Alle drei Anbieter wurden mit den konkreten Anforderungen von BioP II konfrontiert sowie zur Angebotsabgabe aufgefordert. Auf dieser Basis wurde eine Festlegung auf die Systeme von Bundesdruckerei/NEC und Dermalog getroffen.

Zusammenfassend wurden also folgende Anbieter und Systeme für BioP II ausgewählt:

Technologie	Anbieter
Gesichtserkennung	Cognitec Systems GmbH
Fingerabdruckerkennung	Bundesdruckerei GmbH und NEC Deutschland GmbH
	Dermalog Identification Systems GmbH
Iriserkennung	SD Industries GmbH

**Tabelle 1: Systemauswahl für BioP II**

<sup>4</sup> Diese Fragestellung wurde teilweise von den nationalen und europäischen Entwicklungen während der bereits laufenden Untersuchungen überholt, da zwischenzeitlich die Entscheidungen, wie im Vorwort benannt, getroffen wurden.

### 3.3 Projektphasen

Das Projekt BioP II gliederte sich im Wesentlichen in vier Phasen:

- **Konzeption und Beschaffung:** In dieser Phase wurde ein Feinkonzept als Grundlage für die Durchführung der Systemtests erstellt. Wesentlicher Bestandteil war die Konzeption einer Testumgebung und eines Testablaufs, die die Erfassung aller notwendigen Ergebnisse in sinnvoller Weise ermöglichen, sowie der Entwurf einer geeigneten Netzwerkarchitektur zur Erfüllung bestehender Sicherheitsanforderungen bei der Verarbeitung personenbezogener Daten. Aus dem Feinkonzept wurden konkrete Pflichtenhefte für die Hersteller abgeleitet. Des Weiteren wurde die Systembeschaffung organisiert.
- **Implementation und Inbetriebnahme:** Während dieser Phase erfolgte seitens der Hersteller die Anpassung bzw. Implementierung der biometrischen Systeme gemäß Pflichtenheft. Durch secunet wurden die zentralen Komponenten (Datenbank, Netzwerkarchitektur etc.) umgesetzt sowie die erforderlichen Schnittstellen und Auswertemechanismen implementiert. Des Weiteren wurde die Testumgebung vorbereitet. Nach dem Aufbau der Systeme und deren Funktionstests am Flughafen Frankfurt am Main erfolgten Einweisungen des Enrolment-Personals sowie die Aufnahme der Referenzdaten für über 2000 Mitarbeiter von Lufthansa, Fraport und Bundesgrenzschutz<sup>5</sup> (BGS). Dazu wurden seitens Lufthansa und Fraport intensive Werbemaßnahmen durchgeführt (Plakatierung, Intranet, Flyer, Newsletter etc.).
- **Feldtest und weiterführende Untersuchungen:** Am Flughafen in Frankfurt am Main wurden die biometrischen Systeme einem Praxistest mit einer großen Benutzergruppe unterzogen. Der Feldtest beinhaltet gemäß [ISO\_BPTR] sowohl eine „scenario evaluation“ als auch eine „technical evaluation“. Feldtestbegleitend wurden Benutzbarkeits- und Akzeptanzuntersuchungen durch das University College London (UCL) durchgeführt. Darüber hinaus erfolgten weiterführende Untersuchungen hinsichtlich der Speicherung biometrischer Merkmale auf einer RFID-Karte und bezüglich der Überwindungssicherheit der Systeme.
- **Auswertung und Dokumentation:** In dieser Phase erfolgte die Auswertung der im Feldtest und in den weiterführenden Untersuchungen gewonnenen Daten. Diese Ergebnisse sind im vorliegenden Abschlussbericht dokumentiert.

### 3.4 Aufbau des Dokuments

Vorliegendes Dokument stellt den öffentlichen Abschlussbericht des Projekts BioP II dar. Ausgangspunkt der Studie ist die Beschreibung der zu untersuchenden Referenzen in Kapitel 4. Hier werden insbesondere die verwendeten Bilddateien als Referenz für Gesichts-, Fingerabdruck- und Iriserkennungssysteme gemäß ICAO erläutert. Dies beinhaltet sowohl deren Generierung als auch deren Bereitstellung für die biometrischen Systeme im CBEFF-Format (vergleiche [CBEFF]).

In Kapitel 5 werden die Komponenten der beteiligten Testsysteme sowie deren Funktionsweise für BioP II erläutert.

Kapitel 6 beinhaltet eine ausführliche Beschreibung des durchgeführten Feldtests bezüglich Testpopulation, Umgebung, Gesamtarchitektur und Testdurchführung, um die erforderliche Transparenz und Nachvollziehbarkeit zu gewährleisten.

Kapitel 7 stellt schließlich die Auswertung der Feldtestergebnisse dar. Dabei werden zunächst wesentliche Aspekte des Auswertungskonzepts erläutert und motiviert.

---

<sup>5</sup> Die Mitarbeiter des BGS wurden in die Fraport Testteilnehmergruppe hinsichtlich Organisation, Betreuung und Teststandorte integriert.

---

Die Ergebnisse der weiterführenden Untersuchungen sind in Kapitel 8 zusammengefasst. Neben verschiedenen technischen Analysen werden hier die Ergebnisse der Benutzbarkeits- und Akzeptanzuntersuchungen vorgestellt.

Die Grundlage zum Vergleich der beteiligten Verfahren und Systeme stellt das Bewertungsschema in Kapitel 9 dar.

Wesentliches Kapitel bezüglich der Ableitung von Konsequenzen aus BioP II stellt Kapitel 10 mit der Interpretation der Ergebnisse dar. Hier werden im Detail die Testergebnisse hinsichtlich der für BioP II relevanten Fragestellungen interpretiert.

## 4 Beschreibung der verwendeten Referenzbasen

Aufgrund des Fokus von BioP II auf Verwendung biometrischer Merkmale im Zusammenhang mit Personaldokumenten dienen als Referenzdaten Bilder der Merkmale der Testpersonen gemäß den Vorgaben der ICAO [ICAO\_BD]. Außerdem werden systemspezifische Templates als Referenzbasis verwendet, die durch ein Live-Enrolment der Testpersonen erzeugt werden. Letzteres stellt die übliche Betriebsart biometrischer Systeme dar. Damit soll die Template-basierte Erkennungsleistung der Leistungsfähigkeit auf Basis der ICAO-Bilder gegenübergestellt werden.

In BioP II werden die Verfahren der Gesichts-, Fingerabdruck- und Iriserkennung untersucht. Für die Gesichts- und Iriserkennung kommen je ein biometrisches System und für die Fingerabdruckerkennung zwei verschiedene biometrische Systeme zum Einsatz.

Abbildung 4 zeigt eine Übersicht der in BioP II verwendeten Referenzbasen.

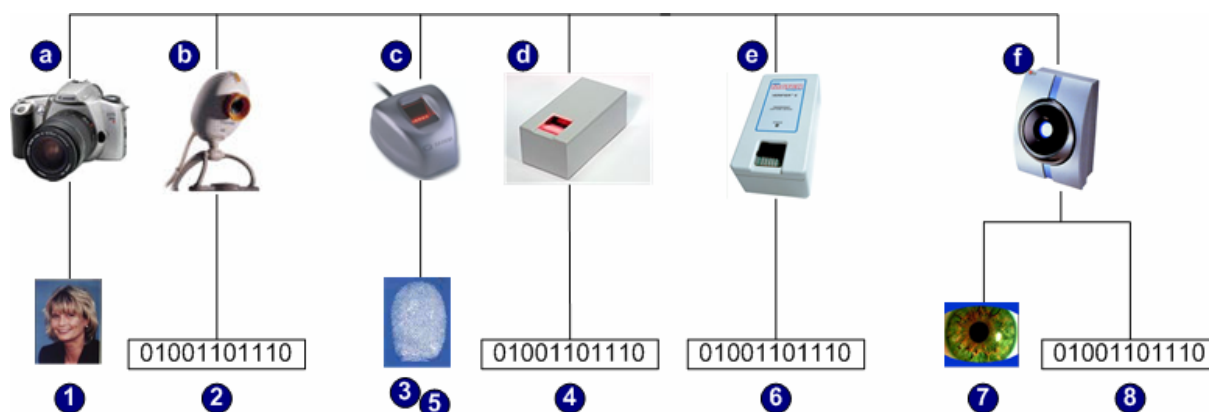


Abbildung 4: Übersicht der BioP II zugrunde liegenden Referenzbasen

RefID	Beschreibung	Bilddatei	Template
1	Bild des Gesichts als Frontalaufnahme	X	
2	System-Template auf Basis Live-Enrolment am GE-System von Cognitec		X
3	Bilder der Fingerabdrücke	X	
4	System-Templates auf Basis Live-Enrolment am FE-System von Bundesdruckerei/NEC		X
5	Bilder der Fingerabdrücke	X	
6	System-Templates auf Basis Live-Enrolment am FE-System von Dermalog		X
7	Bilder der Iriden	X	
8	System-Templates auf Basis Live-Enrolment am IE-System von SD Industries		X

Tabelle 2: In BioP II untersuchte Referenzbasen

Bei der Erzeugung der Referenzbasen kamen folgende Sensoren zum Einsatz (vergleiche Abbildung 4):

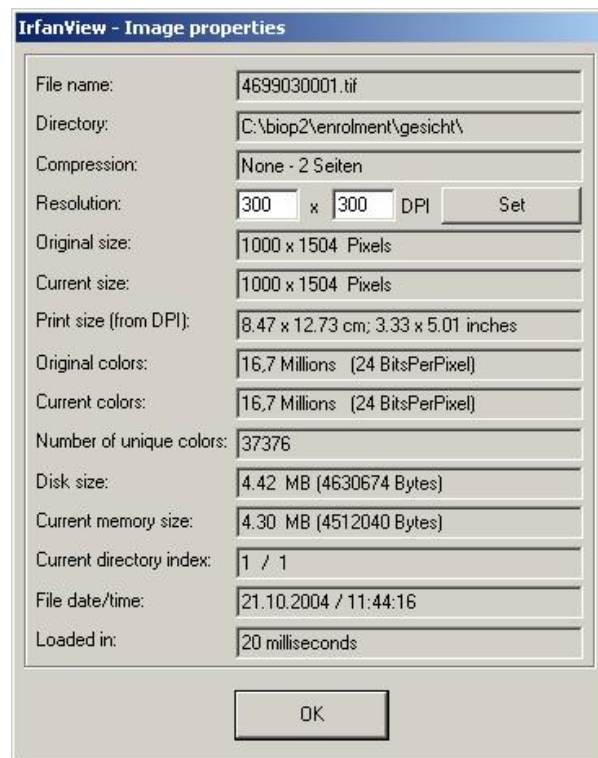
- a) Nikon D100 Digital-Spiegelreflexkamera
- b) Philips ToUcam Pro, PCVC 740K
- c) Sagem MorphoSmart™ MSO 100
- d) Smiths Heimann Biometrics ACCO 1394
- e) CrossMatch Verifier 300
- f) LG Electronics IrisAccess 3000

Die Auswahl der Sensoren erfolgte für b), d), e) und f) durch die am Projekt beteiligten Hersteller, die jeweils ein Komplettsystem inklusive der genannten Sensoren bereitstellten. Die Digital-Spiegelreflexkamera a) wurde durch das BKA empfohlen. Die Festlegung auf c) erfolgte auf Basis der Vorstudie zu BioP II sowie auf Erkenntnissen des BSI zu Fingerabdruckerkennungsverfahren in entsprechenden Projekten. Um keines der am Test beteiligten FE-Systeme zu bevorzugen, war für die Aufnahme der Fingerabdruckbilder der Einsatz eines weiteren, von den beiden FE-Systemen unabhängigen Sensors erforderlich.

#### **4.1 Erstellung von Bilddateien als Referenzbasis**

##### **4.1.1 Erstellung der Gesichtsbilder (RefID1)**

Die Erstellung der Gesichtsbilder erfolgte auf Basis der Richtlinien der ICAO gemäß [ISO\_FACE]. Die Testteilnehmer wurden durch speziell eingewiesene Mitarbeiter des BKA fotografiert. Die Aufnahmen erfolgten in separaten Räumlichkeiten bzw. abgetrennten Bereichen bei Lufthansa und Fraport (vergleiche Abschnitt 6.2.1.1). Als Kamera kam eine Nikon D100 Digital-Spiegelreflexkamera (Objektiv: Nikkor 75 - 300 mm; Blende 4.5 - 5.6) zum Einsatz, mit der die Ausgangsbilder mit den Eigenschaften erzeugt wurden, die in Abbildung 5 wiedergegeben werden.



**Abbildung 5: Eigenschaften eines Ausgangsbildes**

Diese Ausgangsbilder wurden anschließend mit der Bildverarbeitungssoftware Adobe Photoshop 6.0 über die in der Abbildung 6 dargestellte Aktionsliste konvertiert und als JPG-Datei (Kompressionsstufe 2, ohne ICC-Profil) abgespeichert.

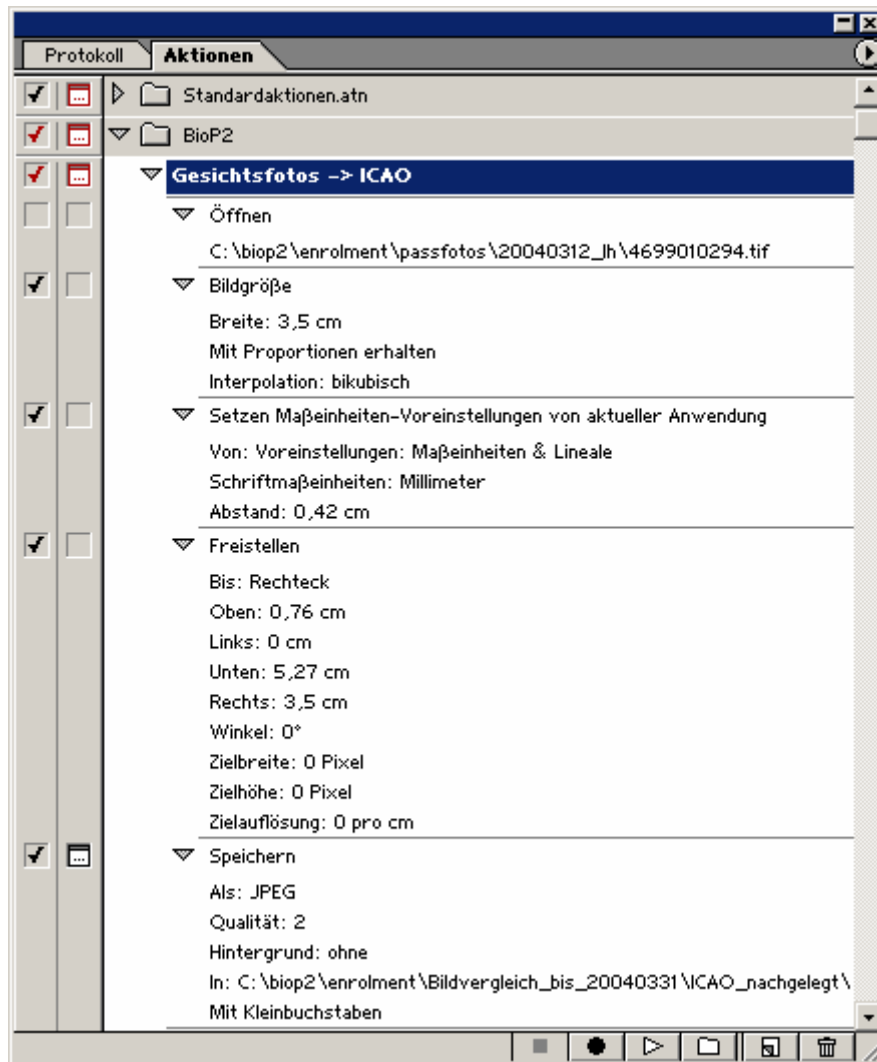
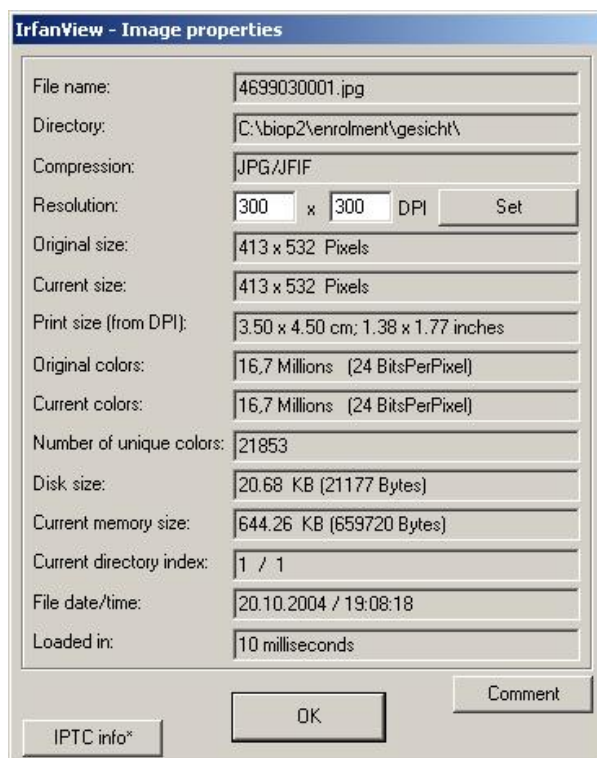


Abbildung 6: Photoshop Aktionsliste für die Generierung der ICAO-konformen Lichtbilder



**Abbildung 7: Exemplarisches ICAO-Gesichtsbild (RefID1)<sup>6</sup>**

In Abbildung 7 ist beispielhaft ein resultierendes ICAO-Gesichtsbild dargestellt. Die Bildeigenschaften sind in Abbildung 8 angegeben.



**Abbildung 8: Eigenschaften eines ICAO-Referenzbildes**

Die finalen Bilddateien resultieren gemäß oben beschriebener Verarbeitung in den in Tabelle 3 dargestellten Dateigrößen. Dort ist außerdem die Dateigröße für die Kodierung im CBEFF-Format entsprechend Abschnitt 4.3 dargestellt.

<sup>6</sup> Die Darstellung aller in diesem Dokument verwendeten Bilder persönlicher Merkmale erfolgt mit Einverständnis der betreffenden Personen.

Dateigröße [kB]	JPG	CBEFF
Minimum	9,7	9,8
Maximum	25,9	25,9
<b>Durchschnitt</b>	<b>13,6</b>	<b>13,7</b>

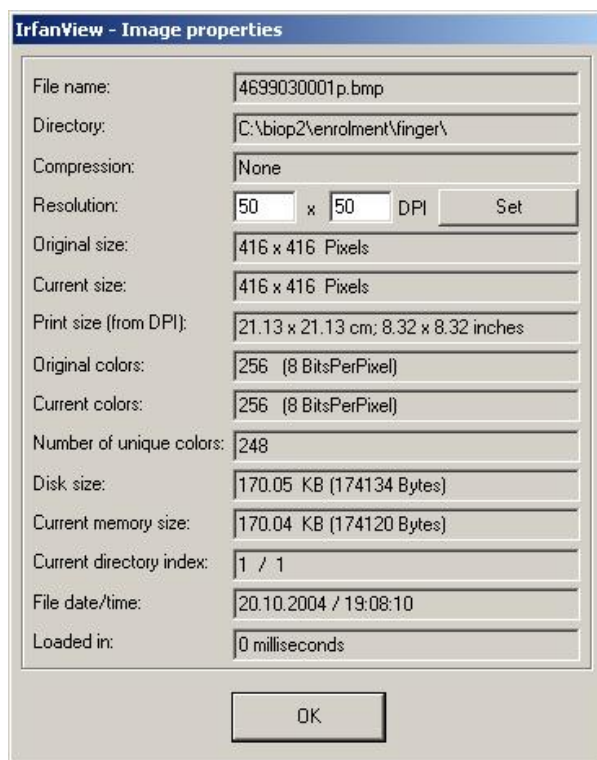
**Tabelle 3: Dateigrößen für RefID1**

#### 4.1.2 Erstellung der Fingerabdruckbilder (RefID3 und RefID5)

Die Aufnahme der Bilder von Fingerabdrücken der Testteilnehmer erfolgte gemäß [ISO\_FINGER]. Als Sensor zur Bildaufnahme kam das System Sagem MorphoSmart™ MSO 100 zum Einsatz. Die Fingerabdrücke wurden flach aufgelegt und nicht gerollt aufgenommen (livescan plain). Die resultierenden Bilder haben eine Größe von 416 x 416 Bildpunkten bei einer Farbtiefe von 8 Bit Graustufen. Ein entsprechendes ICAO-Bild ist beispielhaft in Abbildung 9 dargestellt. Die Bildeigenschaften des zugehörigen Ausgangsbildes im BMP-Format sind in Abbildung 10 angegeben.



**Abbildung 9: Exemplarisches ICAO-Bild eines Fingerabdrucks (RefID3/5)**



**Abbildung 10: Eigenschaften des Ausgangsbildes im BMP-Format**

Als Kompressionsverfahren wurde WSQ<sup>7</sup> angewandt, um eine möglichst starke Kompression zu erhalten und gleichzeitig eine geeignete Qualität zu gewährleisten. Gemäß [ISO\_FINGER] wird bei WSQ eine Kompression bis 15:1 empfohlen, während bei JPG maximal 5:1 vertretbar ist. Die Kompression erfolgte auf Basis einer NIST-Implementierung von WSQ mit der Kompressionsstufe 0,8. Dies entspricht einem Kompressionsfaktor von ca. 16:1 bezogen auf die resultierende Dateigröße.<sup>8</sup>

Die Richtlinien zur Erzeugung und Kompression der Fingerabdruckbilder wurden sowohl mit Daktyloskopen des BKA als auch den beteiligten Herstellern bezüglich Qualität und Eignung zur Erkennung diskutiert und als geeignet betrachtet.

Die finalen Bilddateien resultieren gemäß oben beschriebener Verarbeitung in den in Tabelle 4 dargestellten Dateigrößen. Dort ist außerdem die Dateigröße für die Kodierung im CBEFF-Format entsprechend Abschnitt 4.3 dargestellt.

<sup>7</sup> WSQ (Wavelet Scalar Quantisation) ist ein auf die Kompression von Fingerabdruckbildern spezialisiertes Verfahren.

<sup>8</sup> Gemäß einer im Vorfeld durchgeführten Stichprobe wurde bei der Kompressionsstufe 0,8 ein durchschnittlicher Kompressionsfaktor von 14:1 erwartet. Für alle im Rahmen des Enrolments erfassten Fingerabdrücke ergab sich jedoch eine durchschnittliche Kompression von 16:1.

Dateigröße [kB]	WSQ	CBEFF <sup>9</sup>
Minimum	7,6	15,5
Maximum	14,2	27,6
<b>Durchschnitt</b>	<b>10,6</b>	<b>21,4</b>

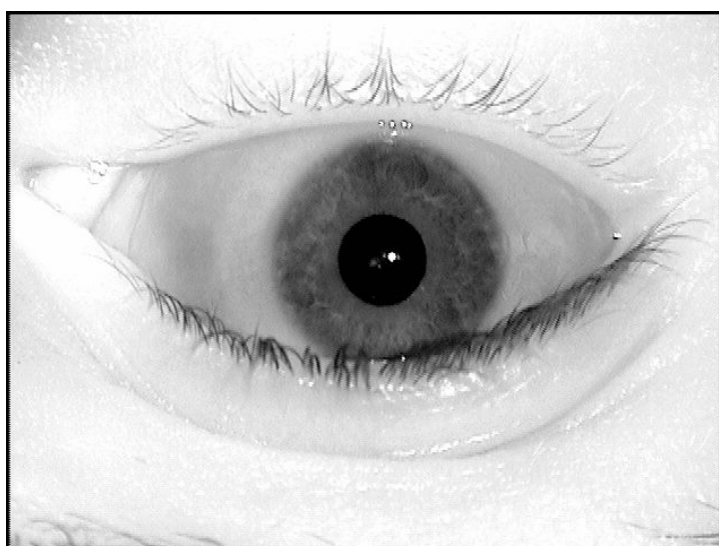
**Tabelle 4: Dateigrößen für RefID3/5**

#### 4.1.3 Erstellung der Irisbilder (RefID7)

ICAO stellt gemäß [ISO\_IRIS] sehr spezielle Anforderungen an die Einheit zur Aufnahme der Irisbilder und die resultierende Bildcharakteristik wie beispielsweise:

- Unterdrückung von Reflexionen durch die Beleuchtung
- Beleuchtung mit Wellenlänge zwischen 700 und 900 nm (Infrarot), Einfallswinkel von mind. 5° zur optischen Achse der Kamera
- Kontrast: mind. 90 Graustufen zwischen Iris und Sklera, mind. 50 Graustufen zwischen Iris und Pupille (zur eindeutigen Bestimmung des inneren und äußeren Randes der Iris)

Eine dedizierte Aufnahmeeinheit zur Erzeugung von Bildern der Iris, die alle ICAO-Anforderungen erfüllt, konnte aufgrund wirtschaftlicher und zeitlicher Restriktionen nicht für die Durchführung des Projekts bereitgestellt werden. Aus diesem Grund wurde für die Aufnahme der Irisbilder die darauf spezialisierte Iriskennungseinheit IrisAccess 3000 von LG eingesetzt, die innerhalb des BioP-Projekts auch im Verifikationssystem von SD Industries zum Einsatz kam. Dieses Gerät ist entsprechend auf die Aufnahme von Irisbildern spezialisiert und berücksichtigt die oben angegebenen ICAO-Anforderungen. Dazu wurden während der Enrolment-Phase Bilder aus den Testverifikationen als BMP-Datei kompressionsfrei gespeichert. Durch das Verwenden der Testverifikation werden die Bilder teilweise durch das Vorhandensein von Brillen gestört. Die Aufnahmen dieser Kamera sind horizontal gespiegelt (vergleiche Abbildung 11).



**Abbildung 11: Exemplarisches ICAO-Bild einer Iris (RefID7)**

<sup>9</sup> Enthält Bilder von zwei Fingerabdrücken (in der Regel rechter und linker Zeigefinger)

Diese Ausgangsbilder wurden anschließend mit der Bildverarbeitungssoftware Adobe Photoshop 6.0 über die in Abbildung 12 dargestellte Aktionsliste konvertiert und als JPG-Datei (Kompressionsstufe 8, ohne ICC-Profil) abgespeichert. Die Festlegung auf Kompressionsstufe 8 erfolgte auf Basis des gemäß [ISO\_IRIS] maximal zulässigen Kompressionsfaktors von 6:1. Die resultierenden Bildeigenschaften sind in Abbildung 13 angegeben.

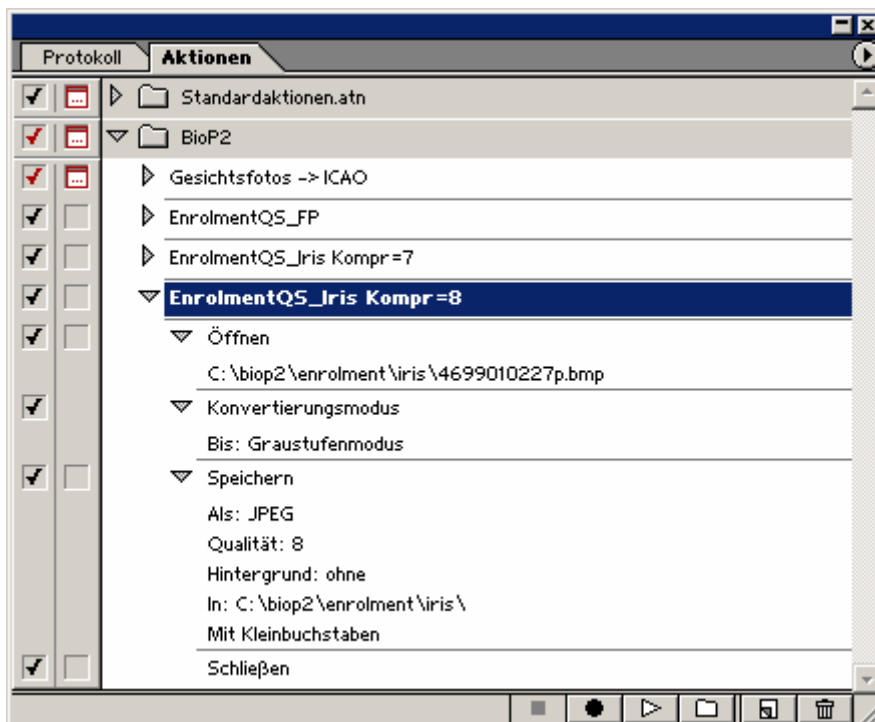
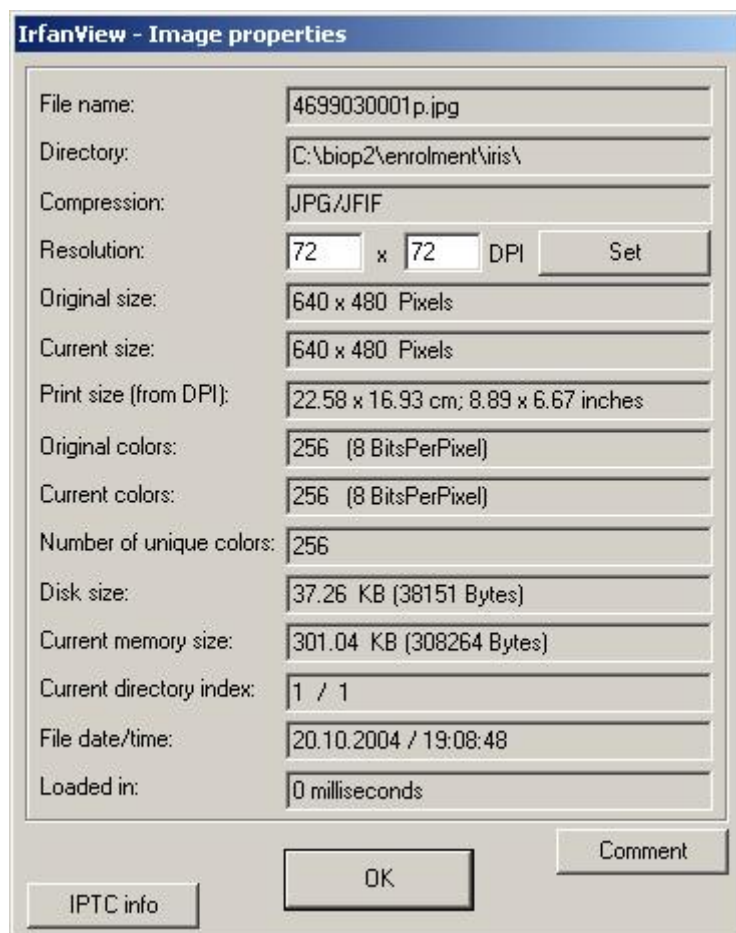


Abbildung 12: Photoshop Aktionsliste für die Konvertierung der Irisbilder



**Abbildung 13: Eigenschaften einer Iris-Bilddatei**

Die finalen Bilddateien resultieren gemäß oben beschriebener Verarbeitung in den in Tabelle 5 dargestellten Dateigrößen. Dort ist außerdem die Dateigröße für die Kodierung im CBEFF-Format entsprechend Abschnitt 4.3 dargestellt.

Dateigröße [kB]	JPG	CBEFF <sup>10</sup>
Minimum	27,4	<b>62,7</b>
Maximum	86,4	<b>167,3</b>
<b>Durchschnitt</b>	<b>50,2</b>	<b>100,6</b>

**Tabelle 5: Dateigrößen für RefID7**

#### **4.2 Erstellung von Templates als Referenzbasis**

Die Templates wurden auf Basis eines Enrolments an den biometrischen Systemen generiert. Die Richtlinien zum Enrolment sind in Abschnitt 6.2.1.8 dargestellt.

<sup>10</sup> Enthält jeweils ein Bild der rechten und linken Iris

### 4.3 Bereitstellung der Referenzdaten für den Test

Die Referenzdaten wurden für den Test über eine zentrale Referenzdatenbank für alle Testorte bereitgestellt. Während die Templates in der durch die Hersteller bereitgestellten Form abgelegt wurden, erfolgte eine Bereitstellung der Bilddateien gemäß ICAO-Vorgaben nicht als Grafikformat, sondern im CBEFF-Format.

#### 4.3.1 Allgemeiner Aufbau der CBEFF-Struktur

Der allgemeine Aufbau der CBEFF-Struktur entspricht dem in [CBEFF] beschriebenen Format.

Standard Biometric Header (SBH)	Biometric Data Block (BDB)	Signature Block (SB)
------------------------------------	-------------------------------	-------------------------

Abbildung 14: CBEFF-Struktur

##### 4.3.1.1 Standard Biometric Header (SBH)

Der CBEFF-Header kann entsprechend der verschiedenen Patron Formate<sup>11</sup> unterschiedlich realisiert werden. Für die Speicherung der biometrischen Bilddaten wird das ISO/IEC JTC 1 SC 37 patron format 3 verwendet (siehe [CBEFF\_Update], Annex D). Dieses Format wurde gewählt, da es die Beschränkung auf die tatsächlich benötigten Elemente ermöglicht und so einen möglichst geringen Overhead für die optionale Speicherung der Daten auf dem RF-Token verursacht.

Innerhalb dieses Headers werden die in Tabelle 6 aufgelisteten CBEFF-Datenelemente verwendet.

CBEFF SBH - gemäß ISO/IEC JTC 1 SC 37 patron format 3 (AllElementsWithDefaultsByteOriented)	
CBEFF_BDB_biometric_organization	Siehe [CBEFF] 6.5.1
CBEFF_BDB_format_type	Siehe [CBEFF] 6.5.2
CBEFF_BDB_security_options	Siehe [CBEFF] 6.5.3
CBEFF_BIR_integrity_options	Siehe [CBEFF] 6.5.4
CBEFF_subheader_count	Siehe [CBEFF] 6.5.5
CBEFF_BDB_biometric_type	Siehe [CBEFF] 6.5.7
CBEFF_BDB_biometric_subtype	Siehe [CBEFF] 6.5.6
CBEFF_version	Siehe [CBEFF] 6.5.22

Tabelle 6: CBEFF-Header

Anders als bei Fingerabdruck- und Irisbildern wird beim Gesichtsbild kein primäres und sekundäres Merkmal (wie zum Beispiel rechter und linker Zeigefinger) abgelegt. Daher wird eine einfache CBEFF-Struktur angewendet, das heißt unmittelbar nach dem ersten CBEFF-Header (SBH) folgt der Block mit dem biometrischen Daten (BDB).

##### 4.3.1.2 Biometric Data Block (BDB)

Der BDB enthält die eigentlichen biometrischen Daten, die durch einen Header gemäß [ISO\_FACE], [ISO\_FINGER] bzw. [ISO\_IRIS] gekapselt werden.

<sup>11</sup> Ein Patron Format ist eine spezifische Definition für CBEFF-Strukturen.

Für die Bilddateien werden im Header die in den nachfolgenden Tabellen beschriebenen Elemente gesetzt.

Facial Image Data – gemäß [ISO_FACE]	
Format Identifier	“FAC”
Version	“010”
Length Of Record	Größe des Elements
Image Data Type	JPEG
Width	Anzahl der Pixel in horizontaler Richtung
Height	Anzahl der Pixel in vertikaler Richtung

**Tabelle 7: Header für Gesichtsbilder**

Finger Image Data – gemäß [ISO_FINGER]	
Format Identifier	“FIR”
Version	“010”
Length Of Record	Größe des Elements
Number Of Fingers	1
Scale Units	Pixel pro Zentimeter oder Pixel pro Inch
Scan Auflösung (X)	Horizontale Scanauflösung
Scan Auflösung (Y)	Vertikale Scanauflösung
Bild Auflösung (X)	Horizontale Bildauflösung
Bild Auflösung (Y)	Vertikale Bildauflösung
Pixel depth	Anzahl Bits pro Pixel
Image Compression Algorithm	WSQ
<i>Angaben pro Finger Image</i>	
Impression Type	Live scan plain
Horizontal line length	Anzahl der Pixel pro horizontale Linie
Vertical line length	Anzahl der horizontalen Linien

**Tabelle 8: Header für Fingerabdruckbilder<sup>12</sup>**

<sup>12</sup> Die Angabe, um welchen Finger es sich handelt, ist im Sub-Header codiert (siehe Abschnitt 4.3.2).

Iris Image Data - gemäß [ISO_IRIS]	
Format Identifier	“IIR”
Version	(Wert in [ISO_IRIS] nicht spezifiziert)
Length Of Record	Größe des Elements
Image Data Type	JPEG

**Tabelle 9: Header für Irisbilder**

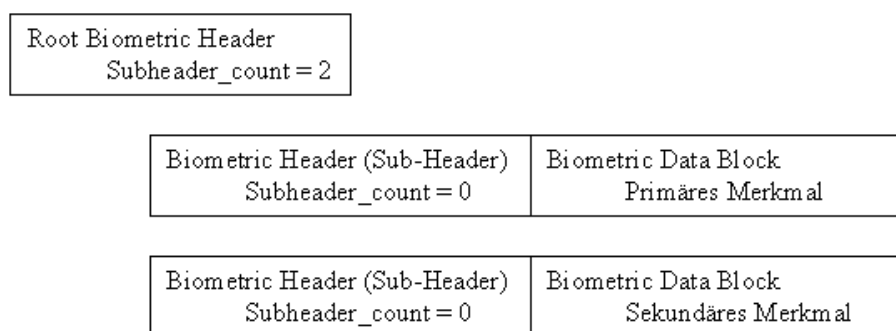
Alle anderen Angaben werden als „unspecified“ angegeben. Im Anschluss an den Header folgen die eigentlichen Bilddaten. Das Gesichtsbild liegt grundsätzlich als Frontalaufnahme (Full Frontal Image) vor.

#### 4.3.1.3 Signature Block (SB)

Der Signature Block kann optional Daten zur Integritätssicherung der CBEFF-Struktur enthalten. In der aktuellen Realisierung wird dieser Block nicht verwendet und ist in dem entsprechenden SBH-Datenelement (CBEFF\_BIR\_integrity\_options) mit „no integrity“ gekennzeichnet.<sup>13</sup>

#### 4.3.2 Speicherung der primären und sekundären Merkmale für Finger und Iris

Die primären und sekundären Merkmale werden innerhalb einer CBEFF-Struktur abgelegt. Hierzu wird das CBEFF-Konzept der geschachtelten BIR-Struktur verwendet.



**Abbildung 15: CBEFF-Struktur zur Speicherung von primärem und sekundärem Merkmal**

Am Anfang der gespeicherten Daten wird ein so genannter „Root Biometric Header“ abgelegt, der durch das Element „CBEFF\_subheader\_count“ angibt, dass zwei weitere CBEFF-Subheader folgen. Jeder dieser Subheader enthält die Informationen zu einem Biometric Data Block, der die eigentlichen biometrischen Daten enthält. Der erste Subheader enthält hierbei die Daten des primären, der zweite Header die Daten des sekundären Merkmals.

Das Verfahren der geschachtelten CBEFF-Header wird bei den Finger-Image-Daten gewählt, obwohl die ISO-Struktur ebenfalls die Speicherung mehrerer Finger-Images innerhalb einer Struktur unterstützt. Durch die Verwendung der CBEFF-Hierarchie ergibt sich eine identische Abbildung bei Iris- und Fingerdaten, was die Weiterverarbeitung der verschiedenen Formate erleichtert.

<sup>13</sup> Die Untersuchung der Sicherheitsmechanismen war nicht Ziel von BioP II. Entsprechende Analysen erfolgen in einem separaten Projekt des BSI (siehe Kapitel 8.3).



## 5 Beschreibung der Systeme

Aufgrund der speziellen Zielsetzung des Projekts BioP II konnten keine Standardsysteme der beteiligten Hersteller zum Einsatz kommen. Stattdessen wurden speziell angepasste Systeme gemäß Pflichtenheften für die Hersteller bereitgestellt, die daher den Status von Prototypen besitzen. Nachfolgend werden die einzelnen Systeme sowie deren Funktionsweise für BioP II beschrieben.

Grundsätzlich setzen sich alle Systeme aus einem Sensor, einer Verarbeitungseinheit, einer Komponente zur Benutzerführung sowie einem kontaktlosen Kartenleser zusammen. Die Erkennungsvorgänge an den Systemen werden durch den Einsatz von kontaktlosen Karten (im Folgenden als RF-Token bezeichnet) ausgelöst. Als RF-Token wird für alle Systeme einheitlich eine Karte des Typs Philips mifare 1k (mifare Standard Card IC MF1 IC S50) eingesetzt.

### 5.1 Systemkomponenten

#### 5.1.1 GE-System von Cognitec

Komponente	Hardware	Software	
		Funktion	Bezeichnung
PC	X86 Comp. Prozessor (VIA/EPIA M10000), 1 GHz 512 MB RAM 18 GB Festplatte	Betriebssystem	Windows 2000 Professional SP3 Version 5.0.2195
		ODBC-Treiber	In der Applikation implementiert
		Remote GUI	TightVNC Server Version 1.2.9
		Kartenleser-Software	rf_api.dll Version 1.1.0.3
		Monitoring Software	Eventsentry Version 2.30
		Gesichtserkennungssoftware	Speziell entwickelte Tools (u. a. BioP II Adapter) auf Basis FaceVACS-Entry 3.3.0
		USV-Systemshutdown	PowerChute Network Shutdown 2.2 build 116
Sensor	Philips ToUcam Pro, PCVC 740K (vierfach), angeschlossen per USB		
RF-Kartenleser	H6156BA mit serielllem Interface; 13,56MHz; ISO 14443A/B; PC/SC		
Feedback	LEDs		

**Tabelle 10: Komponenten des GE-Systems von Cognitec**

### 5.1.2 FE-System von Bundesdruckerei/NEC

Komponente	Hardware	Software	
		Funktion	Bezeichnung
PC	Intel Pentium 4, 2 GHz 512 MB RAM 40 GB Festplatte	Betriebssystem	Windows 2000 Professional SP3 Version 5.0.2195
		ODBC-Treiber	PostgreSQL
		Remote GUI	TightVNC Server Version 1.2.9
		Kartenleser-Software	rf_api.dll Version 1.1.0.3
		Monitoring Software	Eventsentry Version 2.30
		Fingerabdruckerkennungssoftware	Speziell entwickelte Tools (u. a. enrolment.exe, verification.exe)
		USV-Systemshutdown	PowerChute Network Shutdown 2.2 build 116
Sensor	Fingerabdruck-Scanner ACCO 1394 von Smiths Heimann Biometrics, angeschlossen über Firewire		
RF-Kartenleser	H6156BA mit seriellem Interface; 13,56MHz; ISO 14443A/B; PC/SC		
Feedback	TFT-Display mit Ampeldialog und Text		

**Tabelle 11: Komponenten des FE-Systems von Bundesdruckerei/NEC**

### 5.1.3 FE-System von Dermalog

Komponente	Hardware	Software	
		Funktion	Bezeichnung
PC	Intel Pentium 4, 2,6 GHz 256 MB RAM 18 GB Festplatte	Betriebssystem	Windows 2000 Professional SP3 Version 5.0.2195
		ODBC-Treiber	PostgreSQL
		Remote GUI	TightVNC Server Version 1.2.9
		Kartenleser Software	rf_api.dll Version 1.1.0.3
		Monitoring Software	Eventsentry Version 2.30
		Fingerabdruckerkennungssoftware	Speziell entwickelte Tools (u. a. Enrolment.exe, Dermalog.BioPII.Verification.exe)
		USV-Systemshutdown	PowerChute Network Shutdown 2.2 build 116
Sensor	Fingerabdruck-Scanner CrossMatch Verifier 300 angeschlossen über Videoleitung		
RF-Kartenleser	H6156BA mit seriellem Interface; 13,56MHz; ISO 14443A/B; PC/SC		
Feedback	LEDs		

**Tabelle 12: Komponenten des FE-Systems von Dermalog**

#### 5.1.4 IE-System von SDI

Komponente	Hardware	Software	
		Funktion	Bezeichnung
PC	Intel Pentium 4, 2,6 GHz 512 MB RAM 40 GB Festplatte	Betriebssystem	Windows 2000 Professional SP3 Version 5.0.2195
		ODBC-Treiber	PostgreSQL
		Remote GUI	TightVNC Server Version 1.2.9
		Kartenleser-Software	rf_api.dll Version 1.1.0.3
		Monitoring Software	Eventsentry Version 2.30
		Iriserkennungssoftware	Speziell entwickelte Tools (u. a. BiopIISdi.exe)
		USV-Systemshutdown	PowerChute Network Shutdown 2.2 build 116
Sensor	Iriserfassungseinheit IrisAccess 3000 von LG Electronics		
RF-Kartenleser	H6156BA mit seriellem Interface; 13,56MHz; ISO 14443A/B; PC/SC		
Feedback	LEDs und akustische Benutzerführung		

**Tabelle 13: Komponenten des IE-Systems von SDI**



Abbildung 16: GE-System von Cognitec



Abbildung 17: IE-System von SD Industries



Abbildung 18: FE-System von BDr/NEC



Abbildung 19: FE-System von Dermalog

## 5.2 Funktionsweise der Systeme

Gemäß den für die Untersuchungen in BioP II benötigten Funktionen stellen die biometrischen Systeme drei Betriebsmodi bereit:

- **Enrolment-Modus:** Beim Enrolment wird aus dem live von der Testperson präsentierten Merkmal (bzw. Merkmalen) jeweils ein systemspezifisches Template gebildet. Im Anschluss erfolgt mindestens eine Testverifikation.
- **Verifikationsmodus:** Bei der Verifikation wird das Merkmal der Testperson aufgenommen und mit den hinterlegten Referenzdaten verglichen.
- **Batch-Modus:** Im Batch-Modus werden anhand der bei den Verifikationen protokollierten Live-Bilder und Templates Cross-Vergleiche mit den Referenzdaten anderer Testteilnehmer vorgenommen.

### 5.2.1 Enrolment-Modus

Alle biometrischen Systeme stellen für das Enrolment eine Applikation mit grafischer Benutzeroberfläche bereit. Im Einzelnen werden folgende Schritte abgearbeitet (siehe auch Abbildung 20):

- Die Enrolment-Applikation ermittelt die UserID aus dem RF-Token.
- Bei Fingerabdruck- und Iriserkennungssystemen: Der Operator gibt in Auswahlfeldern an, ob das primäre oder das sekundäre Merkmal enroled wird und welcher Finger bzw. welche Iris dies ist.
- Aufnahme primäres Merkmal:
  - **Cognitec:** Von jeder der vier Kameras werden vier Bilder aufgenommen. Das System trifft eine Vorauswahl, welche Bilder für das Enrolment geeignet erscheinen. Der Operator hat die Möglichkeit, gemäß einer Enrolment-Policy einzelne Bilder für die Template-Bildung auszuschließen.
  - **Bundesdruckerei/NEC:** Vom primären Merkmal werden drei Bilder aufgenommen. Die Bildaufnahme erfolgt automatisch, wenn ein Finger auf den Sensor aufgelegt wird. Bilder mit zu schlechter Qualität (zum Beispiel zu wenig Minuten im Bild) werden automatisch ausgeschlossen. Es erfolgt ein Konsistenzcheck, ob alle drei Aufnahmen vom gleichen Finger sind.
  - **Dermalog:** Vom primären Merkmal werden drei Bilder aufgenommen. Die Bildaufnahme wird vom Operator ausgelöst. Es erfolgt ein Konsistenzcheck, ob alle drei Aufnahmen vom gleichen Finger sind.
  - **SD Industries:** Vom primären Merkmal werden in einer festen Zeitspanne fortlaufend Bilder aufgenommen bis eines mit hinreichender Güte erfasst wurde.
- Es wird ein Template aus den aufgenommenen Bildern generiert.
  - **Cognitec:** In die Template-Generierung gehen alle vom System bzw. Operator ausgewählten Bilder ein (Mehrbild-Enrolment).
  - **Bundesdruckerei/NEC:** Von allen drei Bildern des Fingers wird jeweils ein Template gebildet. Anhand eines systeminternen Algorithmus wird das Beste dieser drei Templates ermittelt, die anderen werden verworfen.
  - **Dermalog:** Von allen drei Bildern des Fingers wird jeweils ein Template gebildet. Anhand eines systeminternen Algorithmus wird das Beste dieser drei Templates ermittelt und weiterverwendet, die anderen werden verworfen.

- 
- **SD Industries:** Aus dem aufgenommenen Bild wird der Iriscode (Template) berechnet.
  - Ein Enrolment-Datensatz mit den jeweils zutreffenden Werten wird in die Datenbank geschrieben.
  - Bei Fingerabdruck- und Iriserkennungssystemen: Analog zum primären Merkmal erfolgt die Aufnahme, die Template-Bildung und die Protokollierung für das sekundäre Merkmal in gleicher Weise.
  - Aus der Enrolment-Applikation wird für jedes Merkmal eine Testverifikation gestartet.
  - Abhängig vom Ergebnis der Merkmalsaufnahme und den Testverifikationen kann der Operator eine erneute Merkmalsaufnahme und/oder Testverifikation veranlassen.



- **Toleranzschwelle:** Grenze für den Matchscore, ab welcher eine Verifikation als erfolgreich gewertet wird.
- **Timeout:** Maximale Zeitdauer, während der die Erfassungseinheit fortlaufend Live-Bilder des Merkmals aufzeichnen kann.

Die Einstellung der konfigurierbaren Parameter (Masterreferenz, Toleranzschwelle, Timeout) der Systeme ist in Abschnitt 6.1.6 beschrieben.

Abbildung 21 zeigt schematisch den Funktionsablauf des Gesichtserkennungssystems im Verifikationsmodus.

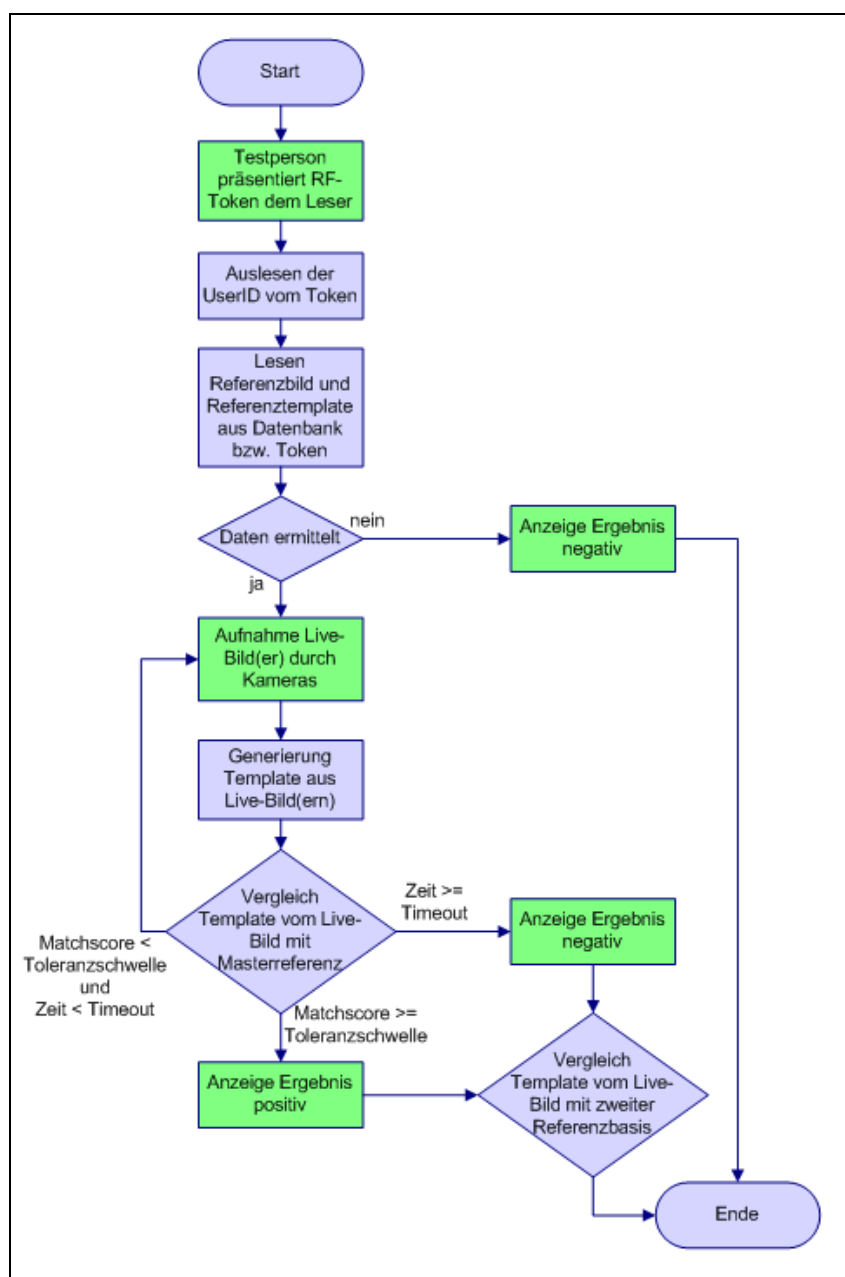


Abbildung 21: Funktionsablauf Gesichtserkennung im Verifikationsmodus

---

Die Testperson präsentiert das RF-Token am RF-Leser des Gesichtserkennungssystems. Nach der Erkennung des Tokens wird die UserID vom System ausgelesen. Anhand der UserID wird der Modus (Referenzdaten in Datenbank oder auf RF-Token) ermittelt und die Daten ausgelesen. Ist dies erfolgreich, werden fortlaufend Bilder des Merkmals aufgenommen und einem Vergleich mit der Masterreferenz unterzogen bis entweder ein positives Verifikationsergebnis vorliegt oder ein voreingestellter Timeout erreicht ist. Das Ergebnis wird mittels LED angezeigt. Abschließend erfolgen eine Verifikation anhand der zweiten Referenzbasis sowie die Protokollierung der Datensätze in die Datenbank.

Abbildung 22 zeigt schematisch den Funktionsablauf der Iris- und Fingerabdruckerkennungssysteme im Verifikationsmodus.

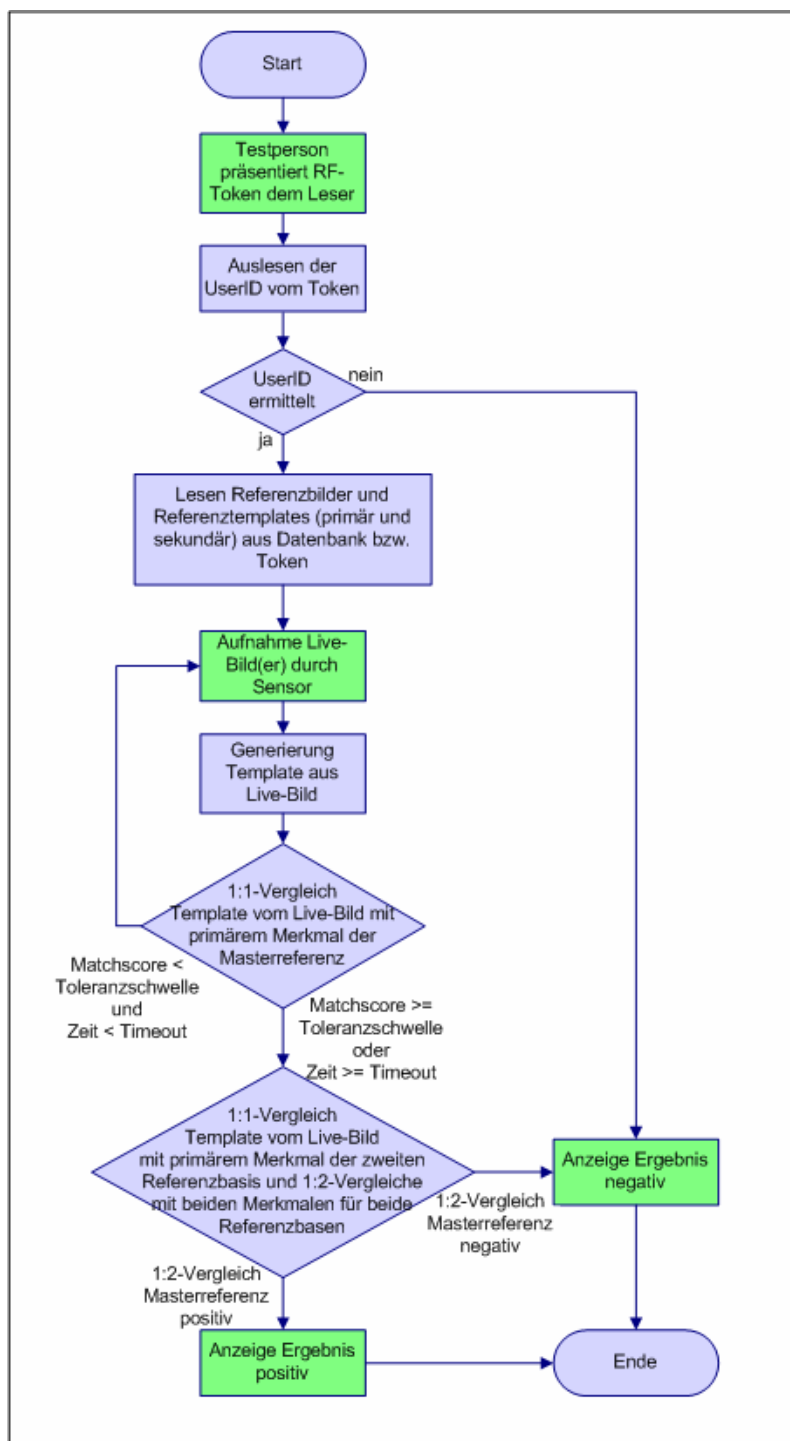


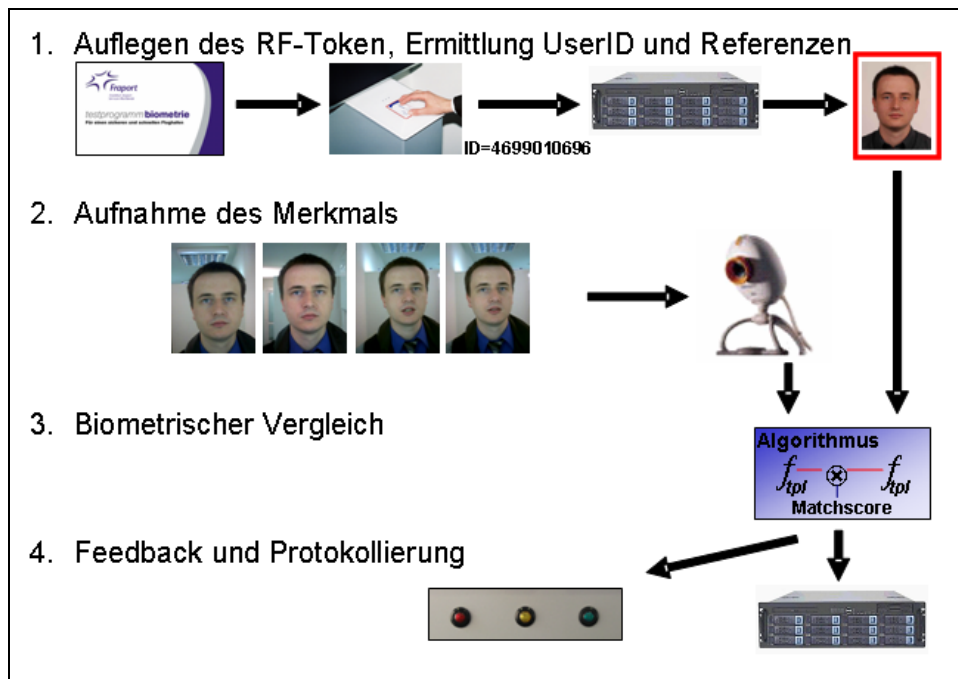
Abbildung 22: Funktionsablauf Iris- und Fingerabdruckerkennung im Verifikationsmodus<sup>14</sup>

Die Funktionsweise ist weitgehend analog zu der Gesichtserkennung. Im Gegensatz zu dieser erfolgen sowohl bei den Fingerabdruck- als auch bei den Iriserkennungssystemen neben den 1:1-Vergleichen mit dem primären Merkmal auch 1:2-Vergleiche mit primärem und sekundärem Merkmal für jede

<sup>14</sup> Bei der Iriserkennung gilt die umgekehrte Relation beim Vergleich von Matchscore mit Toleranzschwelle.

Referenzbasis. Als Feedback wird jeweils das Ergebnis des 1:2-Vergleichs bezüglich der Masterreferenz angezeigt.

Aus Sicht der Benutzer stellt sich der Ablauf vereinfacht dar (siehe Abbildung 23).



**Abbildung 23: Vereinfachter Ablauf einer Betätigung (am Beispiel Gesichtserkennung)**

Das Feedback erfolgt mittels einer Ampel (LED bzw. Bildschirm).<sup>15</sup> Während des Lesens des RF-Token und dem Bezug der Referenzdaten aus der Datenbank blinkt die Ampel gelb, während der Aufnahme des biometrischen Merkmals leuchtet die gelbe Ampel permanent. Bei einer erfolgreichen Verifikation leuchtet die grüne Ampel permanent für etwa drei Sekunden, bei fehlgeschlagener Verifikation leuchtet die rote Ampel permanent für etwa drei Sekunden. Tritt während der Betätigung ein Fehler auf (zum Beispiel Lesen des Token fehlgeschlagen), blinkt die rote Ampel dreimal.

### 5.2.3 Batch-Modus

Für den Batch-Modus (Offline-Verifikationen für die Simulation von Betätigungen Unberechtigter) wurden von den Herstellern separate Tools zur Verfügung gestellt.

Als Input wird jeweils ein Verzeichnis mit Bilddateien oder Templates sowie eine Textdatei mit UserIDs erwartet. Für alle in der Datei angegebenen UserIDs werden jeweils die Referenzdaten aus der Datenbank gelesen und einem Vergleich mit sämtlichen Bildern bzw. Templates in dem Input-Verzeichnis unterzogen. Die Verifikationsergebnisse werden entsprechend in der Datenbank protokolliert.

<sup>15</sup> Das FE-System von BDr/NEC gibt zusätzliches Feedback über eine Textzeile. Das IE-System von SDI verwendet zusätzlich zur Ampel ein akustisches Feedback.

## 6 Feldtestbeschreibung

### 6.1 Testbedingungen

#### 6.1.1 Testpopulation

##### 6.1.1.1 Gesamtpopulation

Zur Bewertung der Zuverlässigkeit der ermittelten Daten sind die Größe der Benutzergruppe und die Anzahl der Versuche pro Testperson entscheidende Parameter. Die Auswahl der konkreten Testpersonen erfolgte durch die beiden beteiligten Unternehmen Lufthansa und Fraport, die jeweils über 1000 Testpersonen stellten.

Insgesamt wurden 2081 Personen enrolt. Von 2025 Teilnehmern wurden demographische Daten erfasst. Die Differenz begründet sich aus Personen, die keine Angaben machen wollten, bzw. Teilnehmern, die nicht alle Stationen des Enrolments beispielsweise aus Zeitgründen absolvierten.

Tabelle 14 zeigt die Zusammensetzung der Testpopulation im Vergleich zur Gesamtbevölkerung der Bundesrepublik Deutschland<sup>16</sup> auf. Die nachfolgenden Diagramme (Abbildung 24 bis Abbildung 28) stellen die Zusammensetzung der Testpopulation grafisch dar.

	Testpopulation		Gesamtbevölkerung		
	Absolut	Relativ [%]	Absolut	Relativ [%]	
<b>Geschlecht</b>	Männlich	1.436	70,91	40.359.000	48,90
	Weiblich	583	28,79	42.172.600	51,10
	Keine Angabe	6	0,30		0
	<b>Gesamt</b>	<b>2.025</b>	<b>100</b>	<b>82.531.600</b>	<b>100</b>
<b>Alter</b>	16-19	29	1,43	3.760.866	5,42
	20-29	375	18,52	9.583.408	13,81
	30-39	657	32,44	12.780.232	18,42
	40-49	582	28,74	12.989.485	18,72
	50-59	345	17,04	9.937.782	14,32
	60 oder älter	29	1,43	20.336.449	29,31
	Keine Angabe	8	0,40		0
	<b>Gesamt</b>	<b>2.025</b>	<b>100</b>	<b>69.388.222</b>	<b>100</b>

<sup>16</sup> Quelle: Statistisches Bundesamt; Stand: 31.12.2003 (Geschlecht und Alter) bzw. Mai 2003 (Ausbildung)

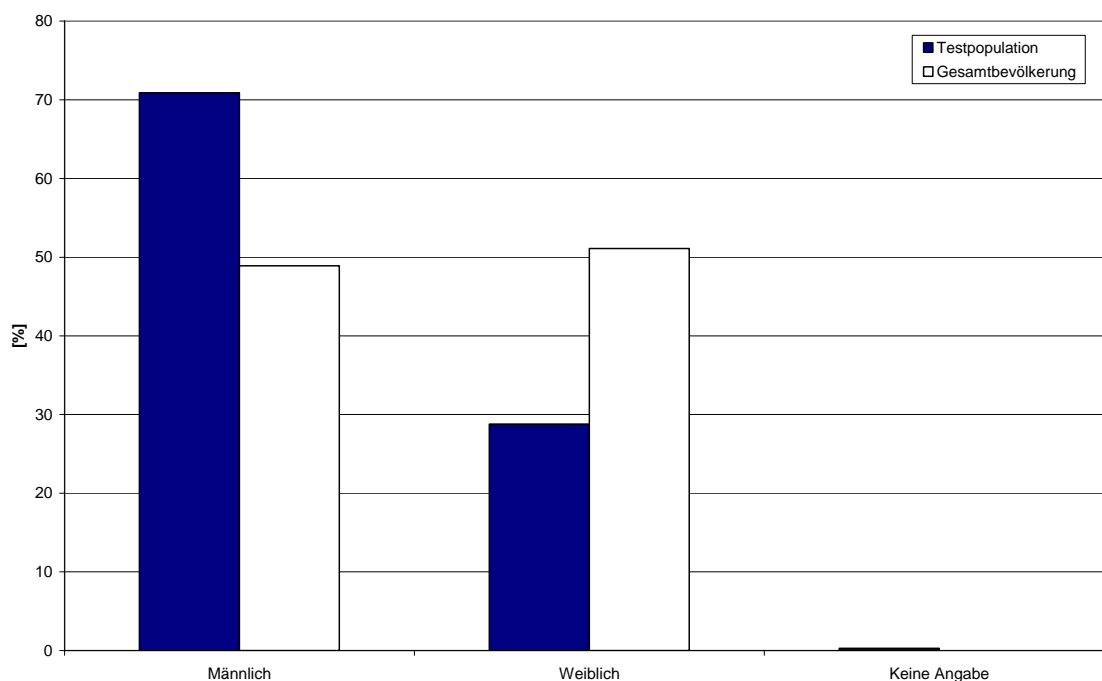
	Testpopulation		Gesamtbevölkerung		
	Absolut	Relativ [%]	Absolut	Relativ [%]	
<b>Ethnische Herkunft</b>	Europäisch	1.977	97,63		
	Afrikanisch	1	0,05		
	Arabisch	4	0,20		
	Asiatisch (indischer Subkont.)	6	0,30		
	Asiatisch (Süd- und Ostasien)	11	0,54		
	Indianisch	2	0,10		
	Andere	16	0,79		
	Keine Angabe	8	0,40		
	<b>Gesamt</b>	<b>2.025</b>	<b>100</b>		
<b>Ausbildung<sup>17</sup></b>	<i>Kaufmännische Ausbildung</i>	778	34,46		
	<i>Handwerklich-technische Ausbildung</i>	581	25,73		
	Ausbildung (Gesamt)	1.359	60,19	38.631.000	58,91
	<i>Akademischer Grad (Natur- oder Ingenieurwissenschaften)</i>	271	12,00		
	<i>Akademischer Grad (Humanwissenschaften, Jura, Management etc.)</i>	317	14,04		
	Akademischer Grad (Gesamt)	588	26,04	7.479.000	11,41
	Andere	191	8,46		0
	Keine Ausbildung	104	4,61	18.423.000	28,09
	Keine Angabe	16	0,71	1.042.000	1,59
	<b>Gesamt</b>	<b>2.258</b>	<b>100</b>	<b>65.575.000</b>	<b>100</b>
<b>Berufliche Tätigkeit<sup>18</sup></b>	Administrativ	1.268	60,67		
	Operativ	716	34,26		
	Andere	89	4,26		
	Keine Angabe	17	0,81		
	<b>Gesamt</b>	<b>2.090</b>	<b>100</b>		

**Tabelle 14: Demographische Daten der Testpopulation im Vergleich zur Gesamtbevölkerung**

Mittels statistischer Testverfahren erfolgte eine Überprüfung, inwieweit die Übereinstimmung der demographischen Strukturen der Testgruppe mit denen der Gesamtbevölkerung der Bundesrepublik Deutschland gegeben ist. Für alle Untersuchungen wurde ein Signifikanzniveau von 0,99 gewählt. Dies bedeutet, dass das verwendete statistische Verfahren mit einer Wahrscheinlichkeit von 99% die richtige Aussage liefert, ob die demographische Struktur übereinstimmt oder nicht.

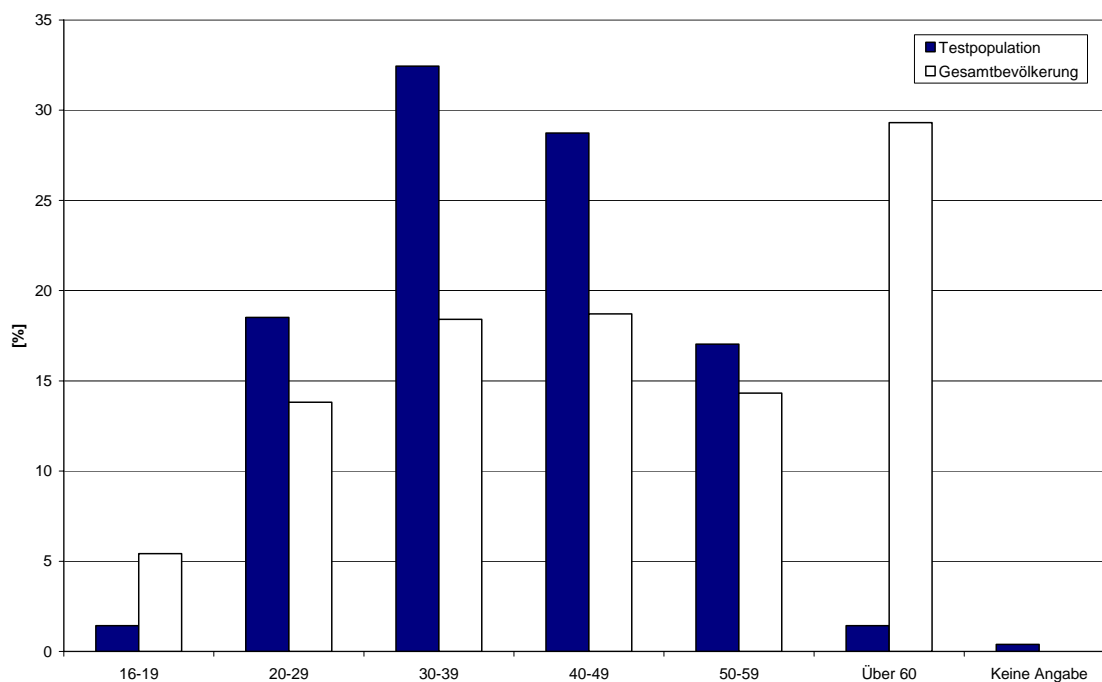
<sup>17</sup> Mehrfachnennungen möglich

<sup>18</sup> Mehrfachnennungen möglich



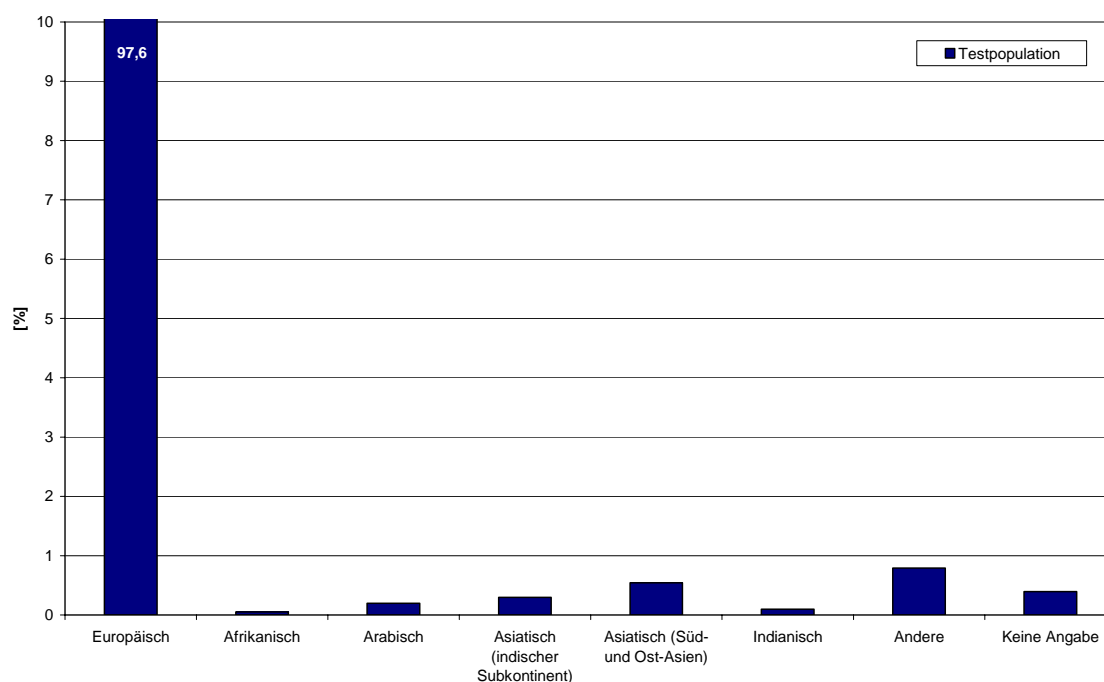
**Abbildung 24: Verteilung der Testpopulation nach Geschlecht**

Hinsichtlich des Merkmals Geschlecht zeigt das statistische Testverfahren eindeutig, dass die Struktur der Testpopulation nicht mit der Struktur der Gesamtbevölkerung übereinstimmt. Dies wird schon daran deutlich, dass das Verhältnis zwischen männlichen und weiblichen Personen in der Testgruppe umgekehrt zur Gesamtbevölkerung ist.



**Abbildung 25: Verteilung der Testpopulation nach Alter**

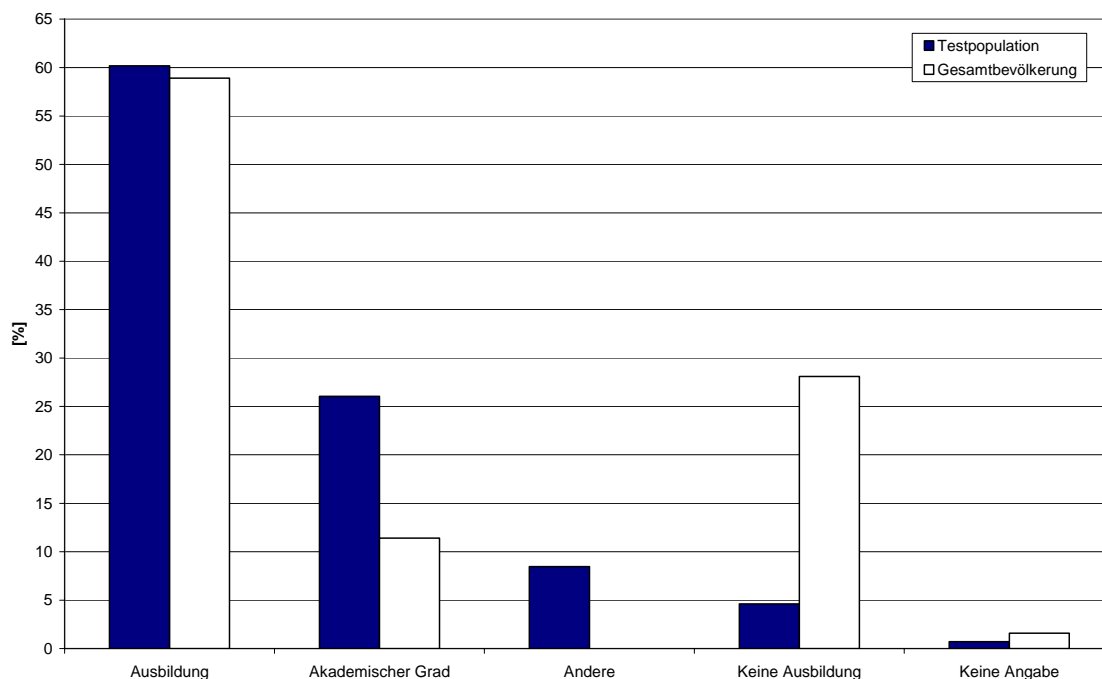
Da sich die Testpersonen aus Mitarbeitern der Fraport AG, der Lufthansa AG sowie des BGS rekrutieren, wurden für den Strukturvergleich sowohl die unter 16-Jährigen als auch die über 65-Jährigen aus der Gesamtpopulation herausgerechnet, da diese das Ergebnis verzerren würden.<sup>19</sup> Trotzdem wird anhand der statistischen Verfahren deutlich, dass die Altersstruktur der Testgruppe von der Gesamtpopulation verschieden ist. Dieses Ergebnis lässt sich dadurch erklären, dass gerade im Randbereich der Altersstruktur sehr große Differenzen zwischen den tatsächlichen und den erwarteten absoluten Häufigkeiten auftreten. Aufgrund der Zusammensetzung der Gesamtbevölkerung hätten weitaus mehr Personen in diese Altersintervalle fallen müssen. Auch bei Einschränkung des Strukturvergleichs auf die Kernbereiche (20-59 Jahre) ist eine Abweichung zwischen der Testpopulation und der Gesamtbevölkerung zu verzeichnen.



**Abbildung 26: Verteilung der Testpopulation nach ethnischer Herkunft**

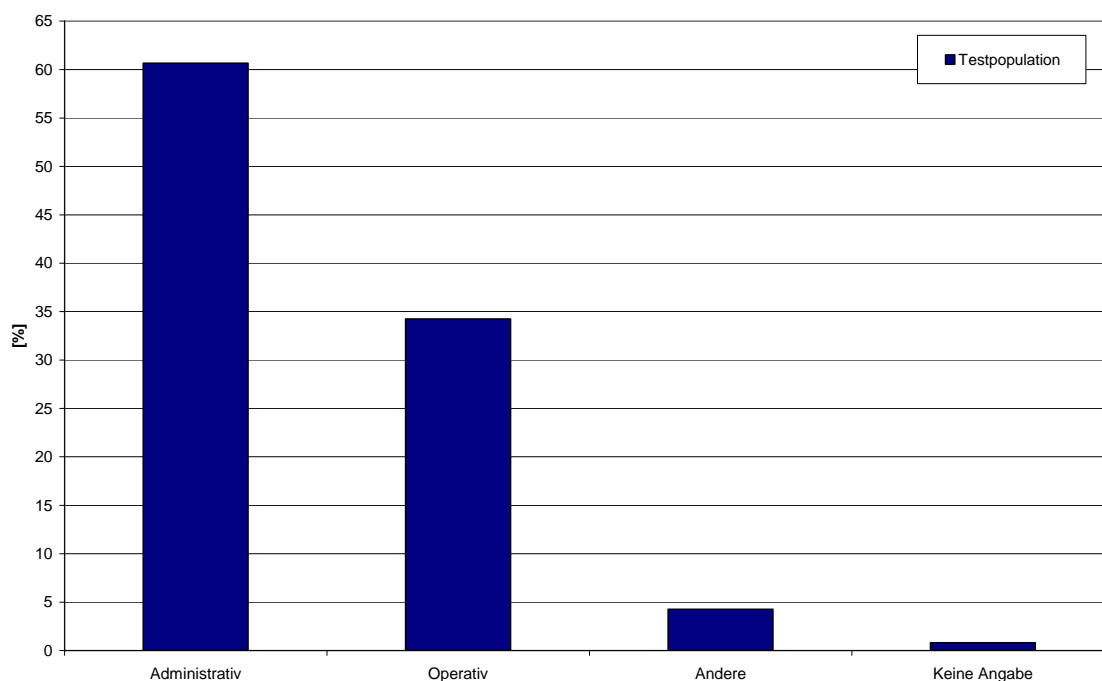
Für die Verteilung der Bevölkerung nach ethnischer Herkunft gibt es keine Daten seitens des Statistischen Bundesamts.

<sup>19</sup> In Abbildung 25 sind auch die über 65-Jährigen mit dargestellt.



**Abbildung 27: Verteilung der Testpopulation nach Ausbildung**

Hinsichtlich der Ausbildung ist die Kategorisierung der Daten des Statistischen Bundesamts leider nicht so ausführlich, wie die Aufteilung der Intervalle bei der Erhebung der Feldtestteilnehmer. Es konnten somit lediglich die Intervalle Lehrausbildung, Akademiker und keine Ausbildung für die Untersuchung verwendet werden. In dem statistischen Test wurden Personen unter 15 Jahre und über 65 Jahre in der Gesamtbevölkerung nicht berücksichtigt, um eventuelle Verzerrungen des Ergebnisses zu vermeiden. Es wird deutlich, dass die Struktur der Ausbildung der Testgruppe nicht gleich der Struktur der Gesamtbevölkerung ist. Aus den Ergebnissen lässt sich ablesen, dass im Bereich der Lehrausbildung keine nennenswerte Differenz in der Struktur besteht. Erwartungsgemäß gibt es weitaus größere Unterschiede in den anderen Bereichen. Zu der Testgruppe gehören mehr Akademiker und weniger ungelernete Arbeiter als nach der Struktur der Gesamtbevölkerung erforderlich wäre.



**Abbildung 28: Verteilung der Testpopulation nach beruflicher Tätigkeit**

Für die Verteilung der Bevölkerung nach beruflicher Tätigkeit gibt es keine vergleichbaren Daten seitens des Statistischen Bundesamts.

Insgesamt lässt sich festhalten, dass die Testpopulation nicht repräsentativ für die Bevölkerung der Bundesrepublik Deutschland ist.

#### 6.1.1.2 Teilpopulationen

Die für die Ergebnisbestimmung relevanten Teilnehmerkreise der einzelnen Systeme weichen von der Gesamtpopulation ab. Dies ist dadurch bedingt, dass verschiedene Teilnehmer keine oder nur sehr wenige Betätigungen durchgeführt haben, andere Teilnehmer hingegen hohe Betätigungszahlen erreichten.

Für jede Systemgruppe wurden gemäß der Betätigungshäufigkeit der Teilnehmer fünf Nutzerklassen nach den in Tabelle 15 dargestellten Kriterien gebildet.

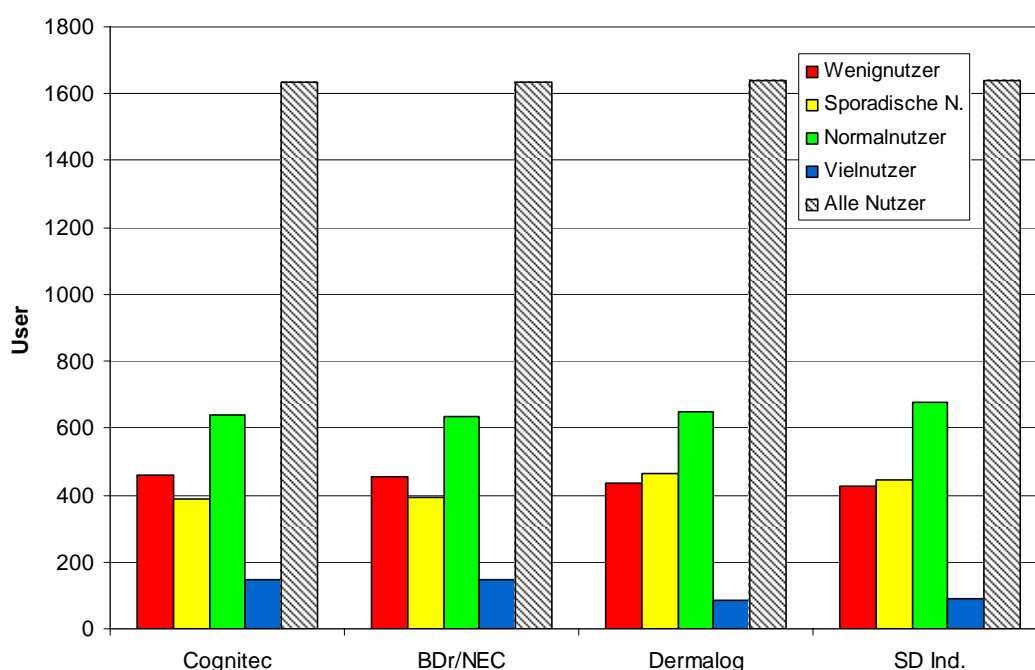
Nutzerklasse	Anzahl auswertbarer Betätigungen für eine Systemgruppe
Wenignutzer	$1 \leq x < 10$
Sporadische Nutzer	$10 \leq x < 30$
Normalnutzer	$30 \leq x < 120$
Vielnutzer	$120 \leq x$
Alle Nutzer	$1 \leq x$

**Tabelle 15: Kriterien zur Bildung systemspezifischer Nutzerklassen**

Der Umfang der einzelnen Nutzerklassen kann Tabelle 16 und Abbildung 29 entnommen werden.

Nutzerklasse	Cognitec	BDr/NEC	Dermalog	SD Industries
Wenignutzer	459	457	437	425
Sporadische Nutzer	389	392	465	445
Normalnutzer	639	636	651	679
Vielnutzer	146	148	85	91
Alle Nutzer	1633	1633	1638	1640

**Tabelle 16: Anzahl Testteilnehmer je Nutzerklasse und System<sup>20</sup>**



**Abbildung 29: Anzahl Testteilnehmer je Nutzerklasse und System**

In Abschnitt 6.1.1.1 wurde dargestellt, dass die Gesamt-Testpopulation nicht repräsentativ für die Bevölkerung der Bundesrepublik Deutschland ist. Für die dargestellten Teilpopulationen können sich Verschiebungen in der demographischen Struktur ergeben. Jedoch ist davon auszugehen, dass auch für die Teilpopulationen keine Repräsentativität für die Gesamtbevölkerung gegeben ist.

### 6.1.2 Teilnehmermotivation

Die Teilnehmer des Feldtests setzten sich aus Mitarbeitern von Lufthansa, Fraport und des BGS zusammen. Die Teilnahme war freiwillig und nicht an etwaige Vergünstigungen (zum Beispiel Arbeitszeitgutschriften) seitens der Arbeitgeber gebunden. Zur Anwerbung der Testteilnehmer wurde das Thema BioP II mittels Plakatierung, Standaktionen, Mitarbeiterzeitschrift und Intranet publiziert.

<sup>20</sup> Die Unterschiede in den Gesamtzahlen von Nutzern für die Systeme resultieren daher, dass verschiedene Teilnehmer nur einzelne Systeme genutzt haben (zum Beispiel aufgrund der Teilnahme an nur einem der beiden Testzyklen).

Im Rahmen des Enrolments konnten die Testpersonen an einer Verlosung teilnehmen, bei der es Uhren mit projektspezifischem Design zu gewinnen gab.

Um einem Sinken der Betätigungszahlen entgegenzuwirken, wurde während des Feldtests eine Incentivierung für Teilnehmer ausgelobt, die eine Mindestanzahl von Betätigungen erreichen.

### 6.1.3 Datenschutz

Gemäß den Anforderungen des betrieblichen Datenschutzes der beteiligten Unternehmen Fraport und Lufthansa wurde ein Sicherheitskonzept erarbeitet und umgesetzt.

Für den Feldtest wurden den Testteilnehmern eindeutige UserIDs zugeordnet. Alle erhobenen Daten, sowohl die bei den Betätigungen angefallenen Protokolldaten als auch die erfragten demographischen Merkmale der Personen, wurden nur anhand der UserID erfasst. Ein direkter Personenbezug konnte nur über Zuordnungslisten hergestellt werden, auf die ausschließlich die beteiligten Unternehmen Lufthansa bzw. Fraport Zugriff hatten.

Alle Testteilnehmer wurden vor Beginn des Feldtests ausführlich über die datenschutzrechtlichen Belange informiert. Dazu wurde unter anderem eine mit dem Bundesbeauftragten für den Datenschutz (BfD) abgestimmte Einwilligungserklärung zur Information der Testteilnehmer verwendet.

### 6.1.4 Testumgebung

Die Durchführung des Feldtests erfolgte an vier verschiedenen Aufstellorten am Flughafen Frankfurt/Main (siehe Abbildung 30). Die Aufstellorte befanden sich im Verantwortungsbereich der Deutschen Lufthansa AG bzw. der Fraport AG:

- Aufstellort „Drehkreuz Ost“ (Lufthansa)
- Aufstellort „Tor 21“ (Lufthansa)
- Aufstellort „Vorfeldkantine“ (Fraport)
- Aufstellort „Vorstandsgebäude“ (Fraport)

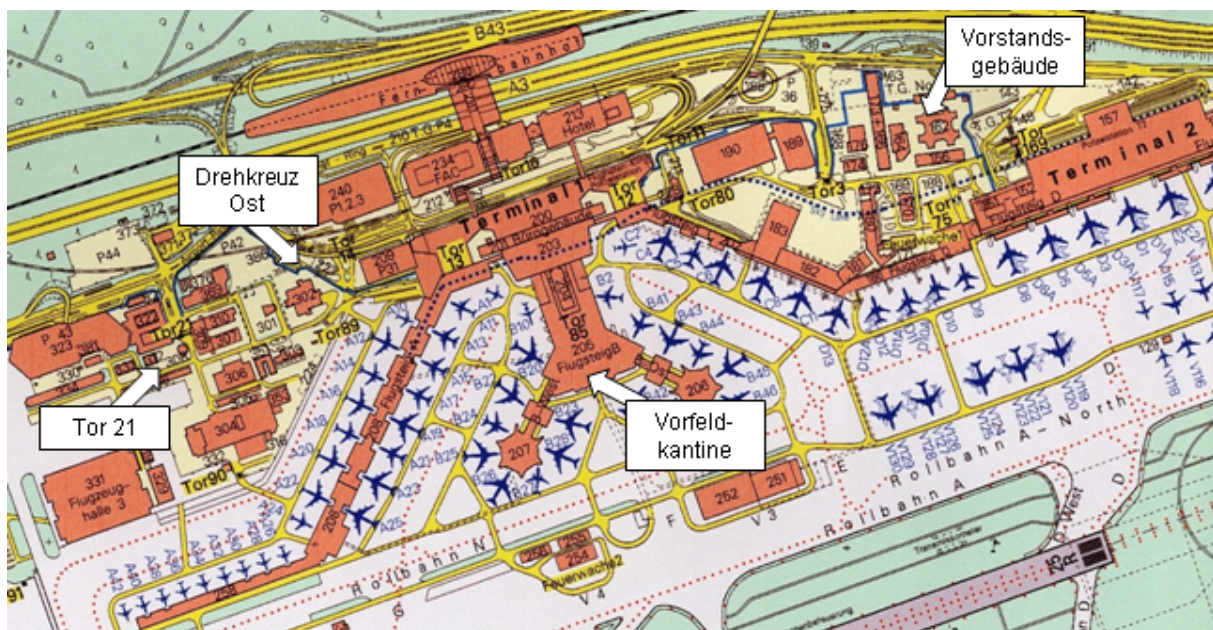


Abbildung 30: Lageplan der Aufstellorte

Mit Ausnahme des Aufstellortes „Vorstandsgebäude“ (siehe Abbildung 31) wurden Container als Testumgebung eingerichtet und aufgestellt. Durch die räumliche Anordnung der biometrischen Systeme in den Containern wurde ein leichtes Durchschleusen der Testpersonen ermöglicht (siehe Abbildung 32 und Abbildung 33).

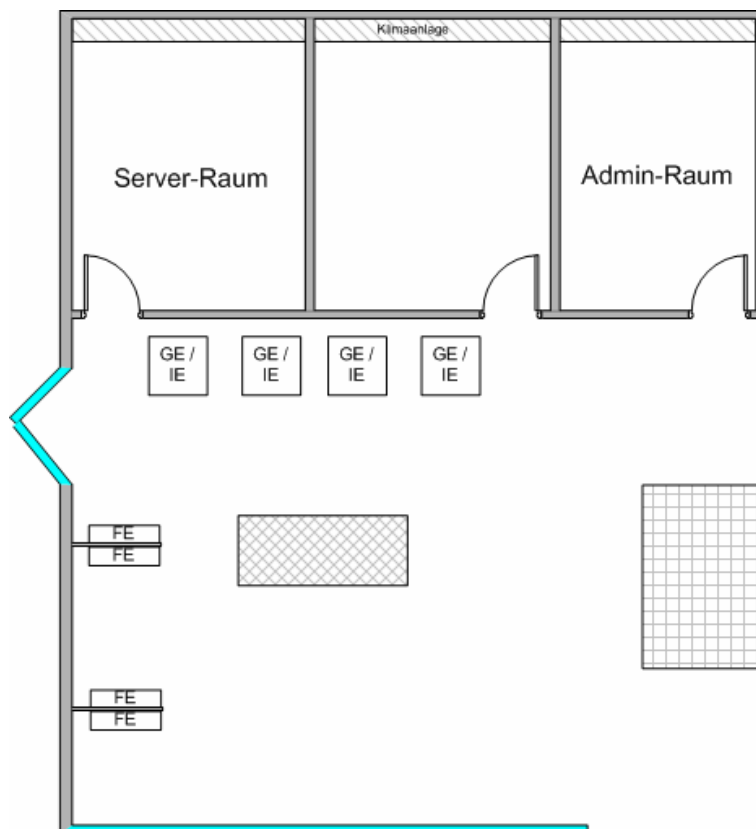


Abbildung 31: Aufstellort Fraport Vorstandsgebäude

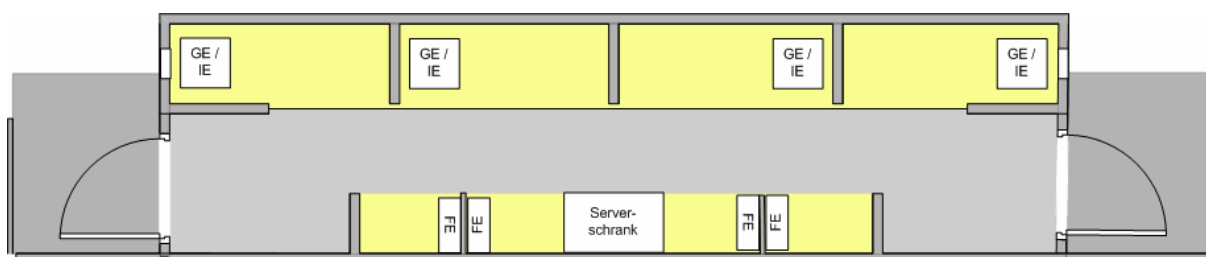
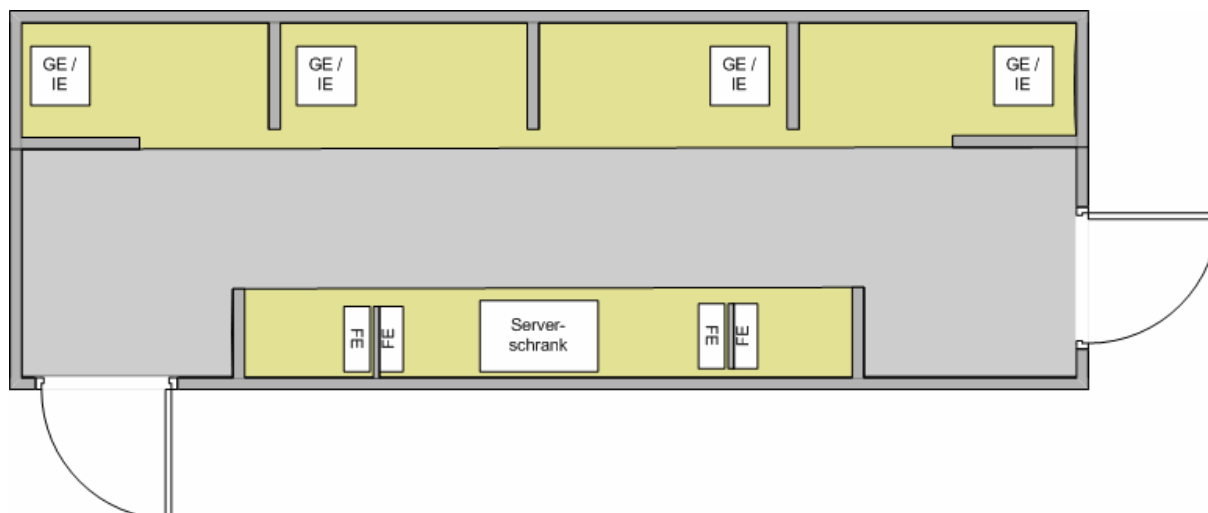


Abbildung 32: Container am Aufstellort Fraport Vorfeldkantine

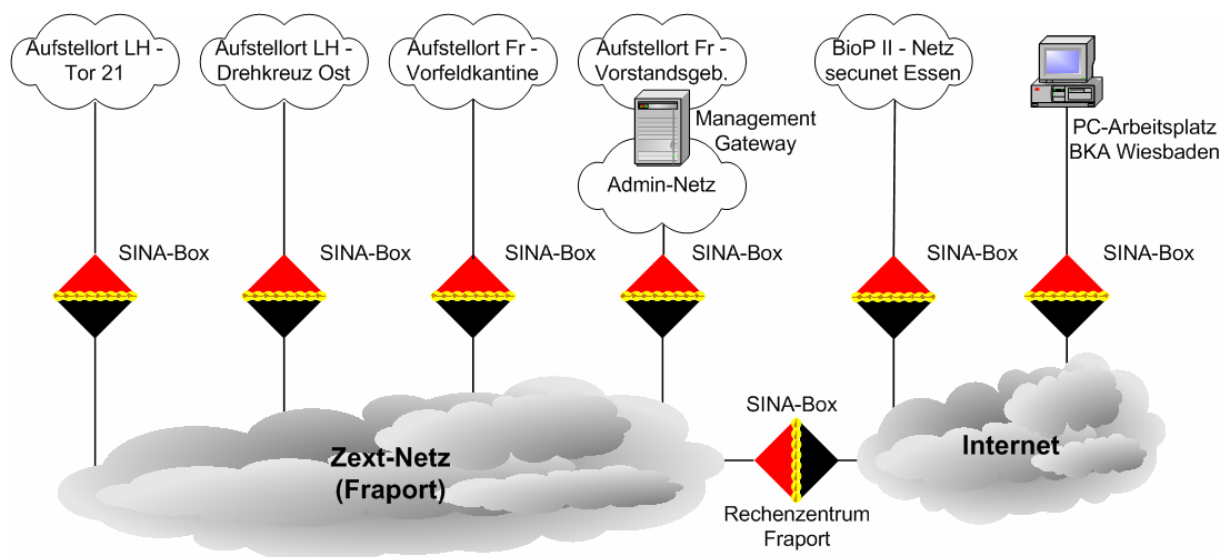


**Abbildung 33: Container an beiden Lufthansa-Aufstellorten**

An jedem Aufstellort wurden jeweils vier biometrische Systeme von zwei Anbietern betrieben. Des Weiteren kam an jedem Aufstellort eine lokale Datenbank zum Einsatz. Diese diente der Bereitstellung der Referenzen für die Systeme. Außerdem wurden sämtliche Ergebnisse in den lokalen Datenbanken protokolliert.

#### 6.1.5 Architektur des Testaufbaus

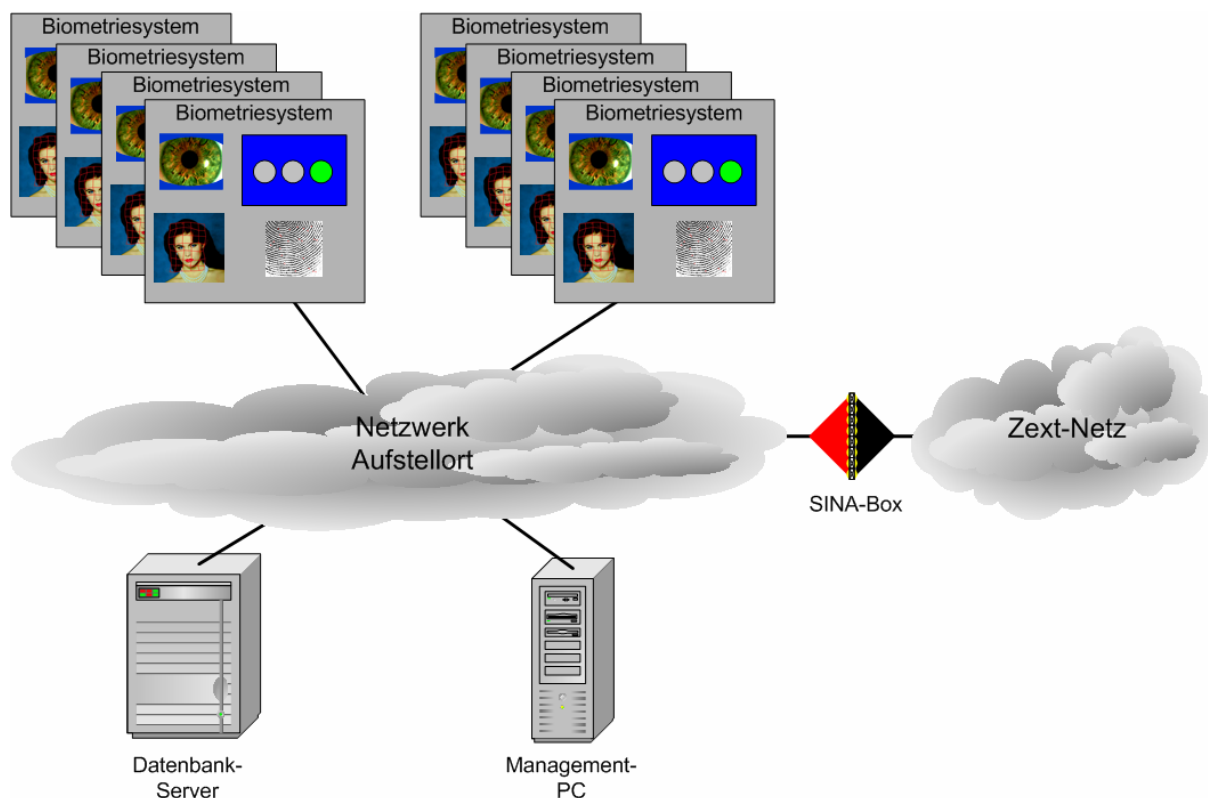
Das BioP-II-Netzwerk war über mehrere Aufstellorte am Flughafen Frankfurt aufgeteilt. Die Kopplung der Aufstellorte erfolgte über das von Fraport bereitgestellte „Zext“-Netz. Die Anbindung des secunet BioP-II-Teilnetzes in Essen und des BKA in Wiesbaden erfolgte über Internetverbindung, die mittels der Hochleistungs-VPN-Lösung SINA des BSI gesichert wurde. Die Anbindung des BioP-II-Netzes an das Internet war durch einen von Fraport bereitgestellten Internetanschluss realisiert. Der Übergabepunkt ins Internet war ebenfalls durch eine SINA-Box abgesichert (siehe Abbildung 34).



**Abbildung 34: Gesamtübersicht BioP-II-Netz**

Die BioP-II-Teilnetze waren als lokal abgeschlossene Netze realisiert. Die standortübergreifende Kommunikation wurde mit SINA gewährleistet. An jedem Aufstellort war eine SINA-Box implementiert, die als VPN-Gateway und Firewall-System arbeitete. Das Risiko einer Kompromittierung der Transportnetze war so sicherheitstechnisch unkritisch. Die räumliche Nähe des Aufstellortes „Fraport Vorstandsgebäude“ mit dem Admin-Netzwerk gestattete es, beide Anbindungen an das Zext-Netz über eine SINA-Box zu koppeln.

An den einzelnen Aufstellorten wurde jeweils ein lokales Netz betrieben, innerhalb dessen die biometrischen Systeme mit der lokalen Datenbank kommunizierten. Zur Administration der Systeme vor Ort enthielt jedes der lokalen Netze einen Management-PC (siehe Abbildung 35).



**Abbildung 35: Infrastruktur Aufstellort**

Im Admin-Netz war die zentrale Datenbank platziert, die ihre Daten aus den einzelnen Datenbanken der Aufstellorte erhielt. Das gesamte Netzwerkmanagement des BioP-II-Netzes erfolgte über Management-PCs, die ebenfalls im Admin-Netz platziert waren.

Das BioP-II-Netz bei secunet in Essen wurde ebenfalls als separates Netz realisiert, welches lediglich über die SINA-Box mit den anderen BioP-II-Teilnetzen gekoppelt war. Innerhalb dieses Netzes waren folgende Systeme enthalten:

- jeweils ein System der beteiligten Biometrie-Hersteller, das mit den im Feldtest eingesetzten Systemen identisch war (für Fehleranalysen sowie weiterführende Untersuchungen)
- lokale Datenbank für die Anbindung der biometrischen Systeme
- zentrale Datenbank als Backup der zentralen Datenbank im Admin-Netz in Frankfurt (Auswertungs-Datenbank)
- Management-Rechner zur Remote-Administration der Systeme in Frankfurt

Ein weiterer Bestandteil des BioP-II-Netzes war ein stand-alone Arbeitsplatz-PC im BKA, der ebenfalls über eine SINA-Box angebunden wurde. Von diesem PC war für die ICAO-konforme Aufbereitung von Referenzdaten sowie für Auswertungszwecke der Zugriff auf einzelne Teile der zentralen Datenbank im BioP-II-Netz bei secunet Essen möglich. Ein Zugang in Teilnetze am Flughafen Frankfurt war seitens des BKA nicht möglich.

### 6.1.6 Systemkonfiguration

Während des Feldtests wurden die biometrischen Systeme ausschließlich im Verifikationsmodus betrieben. Entsprechend musste bei einer Betätigung immer die UserID der betroffenen Person mitgeteilt werden. Die Mitteilung der UserID an das biometrische System erfolgt mittels RF-Token.

Nachfolgend wird die Konfiguration der biometrischen Systeme für den Feldtest dargestellt. Die Parameter sind jeweils so gewählt, dass die Testteilnehmer bezüglich Dauer, Feedback und Erkennungsleistung ein vergleichbares Verhalten der Systeme wahrnehmen, sofern es die verfahrensspezifischen Besonderheiten zulassen. Dies betrifft insbesondere die Parameter Timeout und Toleranzschwelle.

Die Parameter wurden vor Beginn des Feldtests festgelegt und während der gesamten Laufzeit nicht geändert. Die einzige wesentliche Ausnahme bildet Cognitec. Hier war zur Mitte des Feldtests eine Umkonfiguration erforderlich (siehe Abschnitt 6.2.2.1).

#### 6.1.6.1 Parameter der GE-Systeme von Cognitec

Folgende Parameter wurden vor Beginn des Feldtests für alle acht Systeme festgelegt und während des Testzyklus 1 konstant gehalten.

- **Masterreferenz:** Als Referenzbasis, mit der das Live-Bild bei der interaktiven Betätigung einer Verifikation unterzogen wird, wurde das beim Live-Enrolment erzeugte System-Template (RefID2) gewählt.
- **Toleranzschwelle:** Für das Benutzerfeedback wurde die Grenze für den Matchscore, ab welcher eine Verifikation als erfolgreich gewertet wird, auf 0,7 festgesetzt (Wertebereich Matchscore: 0 bis 1, wobei 0 der niedrigste und 1 der höchste theoretisch mögliche Übereinstimmungswert ist).
- **Timeout bei Aufnahme des Live-Bildes:** Die maximale Zeitdauer, während der die Erfassungseinheit fortlaufend Live-Bilder aufzeichnet, wurde auf acht Sekunden gesetzt. Während dieser Zeitspanne konnten vom System in der Regel bis zu fünf Iterationen der Bildaufnahme durchgeführt werden.

Nach der Umkonfiguration zwischen den beiden Testzyklen (vergleiche Abschnitt 6.2.2.1) wurden für alle acht Systeme im Testzyklus 2 folgende Parameter verwendet:

- **Masterreferenz:** Als Masterreferenz wurde das ICAO-Bild (RefID1) gewählt.
- **Toleranzschwelle:** Die Toleranzschwelle wurde auf 0,5 festgesetzt.

Neben den als konfigurierbar geforderten Parametern wurden in Abstimmung mit dem Anbieter folgende Einstellungen vorgenommen:

- **Qualität des aufgenommenen Gesichtsbildes (Gesichtsfinder / Augenfinder):** In mindestens einem der Kamerabilder muss ein Gesicht gefunden werden (AllowedImagesWithoutFace = 3). Es wird nicht vorausgesetzt, dass der Augenfinder ein Ergebnis liefert (AllowedFacesWithoutEyes = 3). Falls der Augenfinder nichts findet, werden

die Augenpositionen vom Gesichtsfinder übernommen. Diese sind jedoch etwas weniger genau.

- **Lebenderkennung:** Die Lebenderkennung ist nicht aktiviert.<sup>21</sup>

#### 6.1.6.2 Parameter der FE-Systeme von Bundesdruckerei/NEC

Folgende Parameter wurden vor Beginn des Feldtests für alle acht Systeme festgelegt und während der beiden Testzyklen konstant gehalten.

- **Masterreferenz:** Als Referenzbasis, gegen die das Live-Bild bei der interaktiven Betätigung einer Verifikation unterzogen wird, wurde das beim Live-Enrolment erzeugte System-Template (RefID4) gewählt.
- **Toleranzschwelle:** Für das Benutzerfeedback wurde die Grenze für den Matchscore, ab welcher eine Verifikation als erfolgreich gewertet wird, auf 800 festgesetzt (Wertebereich Matchscore: 0 bis 20000, wobei 0 der niedrigste und 20000 der höchste theoretisch mögliche Übereinstimmungswert ist).
- **Timeout bei Aufnahme des Live-Bildes:** Die maximale Zeitdauer, während der die Erfassungseinheit fortlaufend Live-Bilder aufzeichnet, wurde auf zehn Sekunden gesetzt. Da für eine Iteration (Bildaufnahme, Merkmalsextraktion, Vergleich) ca. 600ms nötig sind, konnten während dieser Zeitspanne vom System bis zu 16 Iterationen durchgeführt werden.

Neben den als konfigurierbar geforderten Parametern wurden in Abstimmung mit dem Anbieter folgende Einstellungen vorgenommen:

- **Qualität des aufgenommenen Fingerabdruckbildes:**

Folgende Parameter waren durch den Hersteller konfiguriert:

- Minimale Core-Qualität = 1 (erlaubte Werte 1, 2, 4, 8)<sup>22</sup>
- Mindestanzahl Minuten = 20
- Core-Radius = 100<sup>23</sup>

- **Lebenderkennung:** Es waren keine Lebenderkennungsmechanismen konfigurierbar.

#### 6.1.6.3 Parameter der FE-Systeme von Dermalog

Folgende Parameter wurden vor Beginn des Feldtests für alle acht Systeme festgelegt und während der beiden Testzyklen konstant gehalten.

- **Masterreferenz:** Als Referenzbasis, mit der das Live-Bild bei der interaktiven Betätigung einer Verifikation unterzogen wird, wurde das beim Live-Enrolment erzeugte System-Template (RefID6) gewählt.
- **Toleranzschwelle:** Für das Benutzerfeedback wurde die Grenze für den Matchscore, ab welcher eine Verifikation als erfolgreich gewertet wird, auf 50 festgesetzt (Wertebereich

---

<sup>21</sup> Gemäß den Erfahrungen aus BioP I bewirkt die Lebenderkennung eine signifikante Erhöhung der Falschrückweisungsrate.

<sup>22</sup> Der Core ist das Zentrum des Fingerabdrucks.

<sup>23</sup> Der Radius gibt an, in welchem Kreis innerhalb des Fingerabdruckbildes sich der Core befinden muss.

Matchscore: 0 bis 100, wobei 0 der niedrigste und 100 der höchste mögliche Übereinstimmungswert ist).

- **Timeout bei Aufnahme des Live-Bildes:** Die maximale Zeitdauer, während der die Erfassungseinheit fortlaufend Live-Bilder aufzeichnet, wurde auf zehn Sekunden gesetzt. Da für eine Iteration (Bildaufnahme, Merkmalsextraktion, Vergleich) ca. 1s nötig ist, konnten während dieser Zeitspanne vom System bis zu zehn Iterationen durchgeführt werden.

Neben den als konfigurierbar geforderten Parametern wurden in Abstimmung mit dem Anbieter folgende Einstellungen vorgenommen:

- **Qualität des aufgenommenen Fingerabdruckbildes:** Der Schwellwert (QualityThreshold) war auf 50 festgelegt (Wertebereich 0 bis 100).
- **Sensorempfindlichkeit:** Im Verifikationsmodus erkennt das System automatisch, ob ein Fingerabdruck auf dem Scanner aufliegt oder nicht („Auto Detection“). Die Qualität der aufgenommenen Fingerabdrücke hängt von der Empfindlichkeit der Auto Detection ab. Ist diese zu empfindlich eingestellt, wird ein Bild des Fingerabdrucks bereits aufgenommen, wenn der Finger noch nicht ganz aufliegt. Bei einer zu unempfindlichen Einstellung dagegen wird unter Umständen nicht erkannt ob ein Finger aufliegt. Die Empfindlichkeit wurde in Abstimmung mit dem Hersteller mehrfach angepasst. Die letztendlich gewählte Konfiguration war V300=213 (ab 12.05.2004).
- **Lebenderkennung:** Es waren keine Lebenderkennungsmechanismen konfigurierbar.

#### 6.1.6.4 Parameter der IE-Systeme von SD Industries

Folgende Parameter wurden vor Beginn des Feldtests für alle acht Systeme festgelegt und während der beiden Testzyklen konstant gehalten.

- **Masterreferenz:** Als Referenzbasis, mit der das Live-Bild bei der interaktiven Betätigung einer Verifikation unterzogen wird, wurde das beim Live-Enrolment erzeugte System-Template (RefID8) gewählt.
- **Toleranzschwelle:** Für das Benutzerfeedback wurde die Grenze für den Matchscore, ab welcher eine Verifikation als erfolgreich gewertet wird, auf 0,33 festgesetzt<sup>24</sup> (Wertebereich Matchscore: 0 bis 1, wobei 0 der höchste und 1 der niedrigste mögliche Übereinstimmungswert ist).
- **Timeout bei Aufnahme des Live-Bildes:** Die maximale Zeitdauer, während der die Erfassungseinheit fortlaufend Live-Bilder aufzeichnet, wurde auf 13 Sekunden gesetzt. Innerhalb dieser Zeit wurden fortlaufend Aufnahmen des Merkmals erzeugt, von denen bis zu drei einer Merkmalsextraktion und Verifikation unterzogen wurden.

Neben den als konfigurierbar geforderten Parametern wurden in Abstimmung mit dem Anbieter folgende Einstellungen vorgenommen:

- **Qualität des aufgenommenen Irisbildes:** Dieser Parameter war nicht konfigurierbar.
- **Lebenderkennung:** Die Lebenderkennung war gemäß Herstellerkonfiguration aktiviert.

---

<sup>24</sup> Dieser Schwellwert ist durch das zugrunde liegende SDK von LG Electronics fest vorgegeben und nicht konfigurierbar.

### 6.1.7 Wartung und Pflege der biometrischen Systeme

Die biometrischen Systeme wurden in einer unüberwachten Umgebung betrieben. Mindestens zweimal täglich erfolgten Probeverifikationen zur Funktionsprüfung der Systeme. Dabei wurden die Fingerabdrucksensoren gereinigt. Darüber hinaus hatten die Testteilnehmer die Möglichkeit, mit an den Standorten bereitstehenden Papiertüchern bei Bedarf selbst die Sensoren zu reinigen.

## 6.2 Testdurchführung

Die Testdurchführung gliederte sich im Wesentlichen in fünf Abschnitte (vergleiche auch Abbildung 36):

1. Anwerbung von Testteilnehmern (16.02.2004 bis 08.04.2004)
2. Enrolment (15.03.2004 bis 08.04.2004)
3. Feldtest Zyklus 1 inkl. TeachIn<sup>25</sup> (19.04.2004 bis 08.06.2004)
4. Tausch der Systeme (Abbau und Neuinstallation) zwischen den Aufstellorten eines Standortes (11.06.2004 bis 13.06.2004)
5. Feldtest Zyklus 2 inkl. TeachIn (14.06.2004 bis 08.08.2004)

Im Folgenden wird die Testdurchführung je Abschnitt detailliert beschrieben.

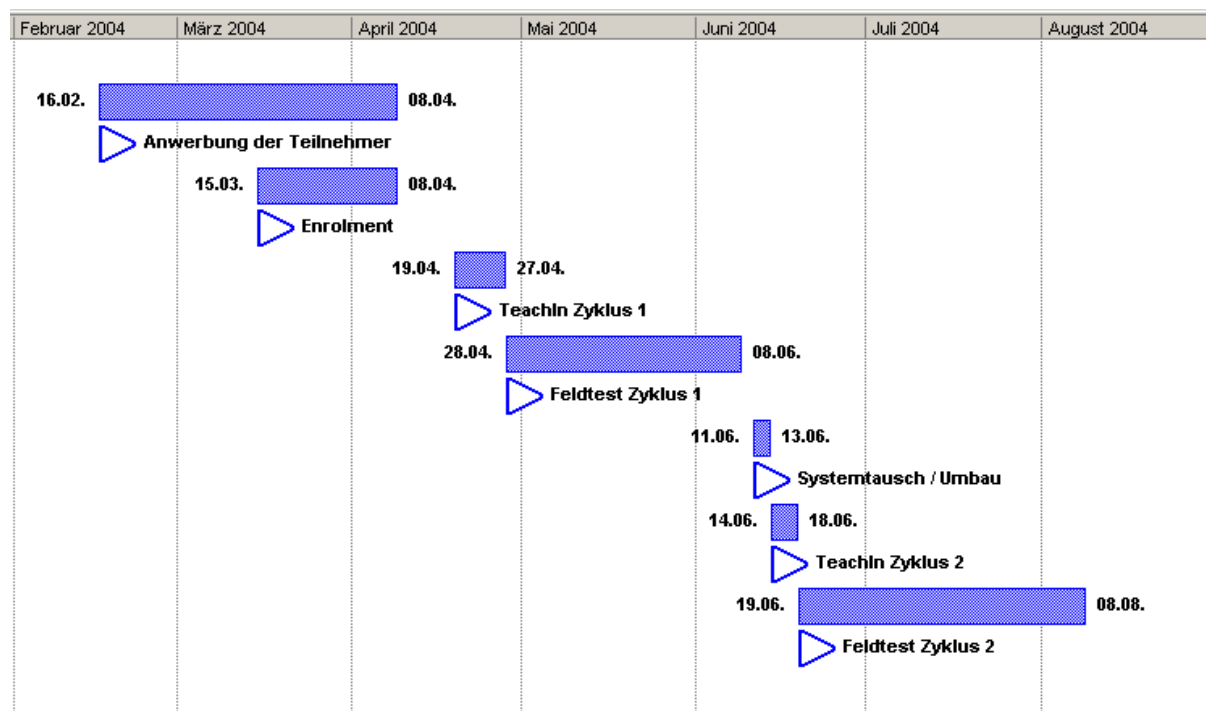


Abbildung 36: Zeitplan Testablauf

### 6.2.1 Anwerbung und Anmeldung der Testpersonen

Die Testdurchführung startete mit der Akquise der Testpersonen. Die Anwerbung bzw. Anmeldung zum Test erfolgte vom 16.02.04 bis zum 08.04.04 für Fraport- und Lufthansa-Mitarbeiter elektronisch

<sup>25</sup> TeachIn bezeichnet eine dem regulären Testzyklus vorgeschaltete Trainingsphase.

über das jeweilige Intranet oder schriftlich über bei Standaktionen verteilte Postkarten. Die Anmeldung der Teilnehmer beim Arbeitgeber (Lufthansa bzw. Fraport) beinhaltete folgende Punkte:

- Einwilligungserklärung
- Registrierung

**Einwilligungserklärung:** Eine Vorlage wurde vom BSI bereitgestellt und mit den betrieblichen Datenschutzbeauftragten und dem Betriebsrat beider beteiligter Unternehmen (Fraport und Lufthansa) abgestimmt. Der Testteilnehmer musste den in der Erklärung dargestellten Punkten verbindlich zustimmen. Dies erfolgte durch schriftliche Unterzeichnung der Erklärung.

**Registrierung:** Die Registrierung beinhaltete die Erfassung von Daten der Testteilnehmer durch den Arbeitgeber. Erfasst wurden:

- Name, Vorname
- Personalnummer
- Erreichbarkeit (Telefonnummer, E-Mail-Adresse, Abteilung etc.)
- Gewünschter Ort und Termin des Enrolments

In Abbildung 37 ist beispielhaft das Anmeldeformular im Intranet eBase der Lufthansa dargestellt.

The screenshot displays a web browser window with the following elements:

- Browser Title Bar:** Testprogramm Biometrie | Projekt Biometrietest | Anmeldung: Schritt 2 - Microsoft Internet Explorer
- Address Bar:** https://www.ebase.dlh.de/home/index.php?n=/home/projekte/biometrie/library/nav3.htm&h=/home/projekte/biometrie/library/hotlinks.htm&c=http://biometrie.ebase.dlh.de/register.php
- Page Header:** Home Deutsch | Home English | Suche [ ] | alle Seiten | Los | Aktien: LHAG 14,27 +0,29
- Navigation:** Arbeitsplatz | Nachrichten | Arbeit & Leben | Wissen | Reisen | Marktplatz
- Main Content:**
  - Willkommen bei Lufthansa**
  - Anmeldung | Schritt 1 > Schritt 2 > Schritt 3**
  - Ihr Name \***
    - Anrede: [Bitte wählen Sie...]
    - Vorname: [ ]
    - Nachname: [ ]
  - Ihre E-Mail-Adresse \***: [ ]
  - Ihre PK-Nummer \*** (z.B. 123456A): [ ]
  - Ihre Abteilung \*** (z.B. FRA XY/Z): [ ]
  - Ihre Telefonnummer**
    - dienstlich \*: [ ]
    - mobil: [ ]
- Buttons:** Zurücksetzen, Weiter
- Footer:** 08.12.03 | FRA ES/X

- Sidebar (Left):**
- Biometrie**
  - Testprogramm
  - Projekt Automatisierte Grenzkontrolle
  - Projekt Biometrietest
    - Mitmachen lohnt
    - Teilnahme
    - Anmeldung
    - Enrolment
    - Projektpartner
    - Datenschutz
    - Häufige Fragen
  - Ansprechpartner
  - FAQs
- Hotlinks**
  - Anmeldung
  - Lageplan
  - Enrolment-Termine
- Sidebar (Right):**
- testprogramm biometrie**
  - Mitmachen lohnt sich!
- Video**
  - „Geben Sie uns eine Minute“ – Der Biometrie-Film mehr...
- Weitere Infos**
  - Biometrische Merkmale
  - Zeitplan Projekt Biometrietest
  - Werbekampagne
  - News-Archiv
  - Bundesgrenzschutz (www)
  - Biometrie in der Airline-Branche (www)

Abbildung 37: Registrierungsmaske der Lufthansa

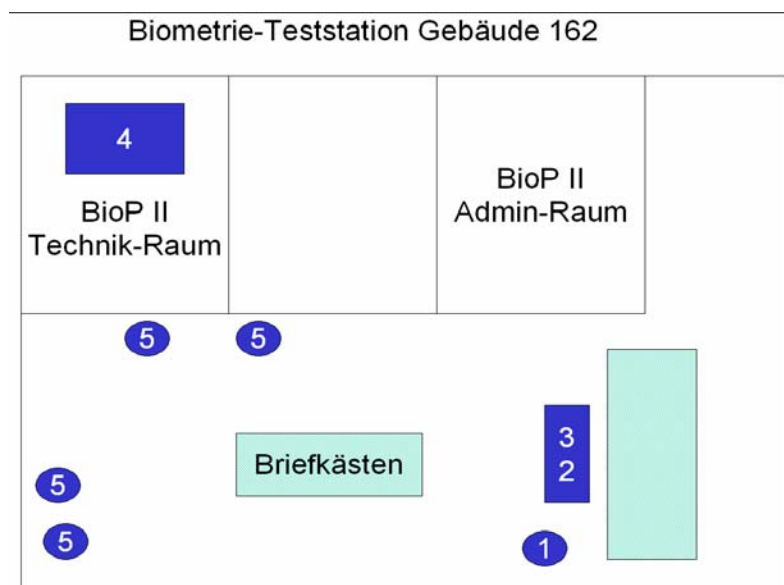
Bis zum 12.03.04, also dem Start des Enrolments, hatten sich mehr als 700 Lufthansa-Mitarbeiter und 490 Fraport-Mitarbeiter angemeldet. Während des Enrolments vom 15.03.04 bis zum 08.04.04 waren in den Enrolment-Stationen so genannte Spontanmeldungen möglich. Hierdurch wurden weitere 900 Teilnehmer gewonnen. Die Testpersonen füllten zur Anmeldung einen vordruckten Anmeldebogen aus, der eine UserID sowie die Einverständniserklärung enthielt.

### 6.2.1.1 Enrolment-Center

Das Enrolment erfolgte in so genannten Enrolment-Centern. Aufgrund der zur Verfügung stehenden Räumlichkeiten sowie technischen und organisatorischen Randbedingungen wurde der Aufbau der Enrolment-Center für Fraport und Lufthansa unterschiedlich ausgelegt. Fraport stellte im Vorstandsgebäude und der Vorfeldkantine zwei Enrolment-Center (vergleiche Abbildung 38 und Abbildung 39), die jeweils die Blöcke eins bis fünf umfassten. Lufthansa stellte im Biometrie-Infocenter ein Enrolment-Center (vergleiche Abbildung 40), welches die Blöcke eins bis vier umfasste und ein weiteres Enrolment-Center (Tor21), das Block fünf umfasste.

Die einzelnen Blöcke beinhalten folgende Aktionen:

- Block 1: Ausgabe der RF-Karte (siehe Abschnitt 6.2.1.4)
- Block 2: Demographische Erfassung (siehe Abschnitt 6.2.1.5)
- Block 3: Bildaufnahme der Fingerabdrücke (siehe Abschnitt 6.2.1.6)
- Block 4: Erstellung der Passbilder (siehe Abschnitt 6.2.1.7)
- Block 5: Enrolment an den biometrischen Systemen (siehe Abschnitt 6.2.1.8)



**Abbildung 38: Enrolment-Center im Fraport Vorstandsgebäude**

Biometrie-Teststation Vorfeldkantine

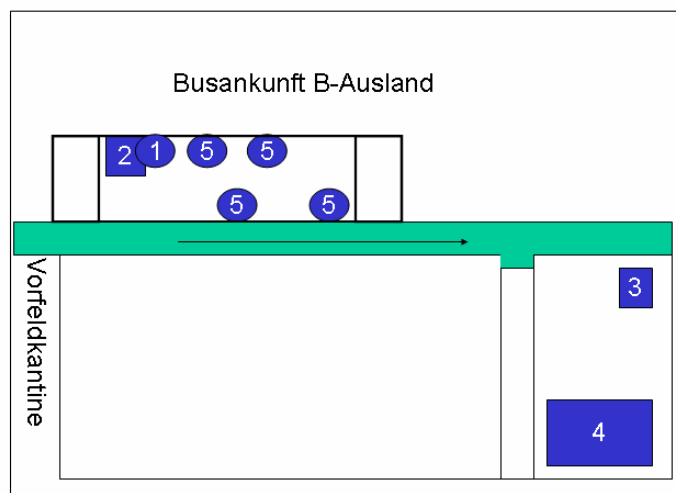


Abbildung 39: Enrolment-Center von Fraport an der Vorfeldkantine

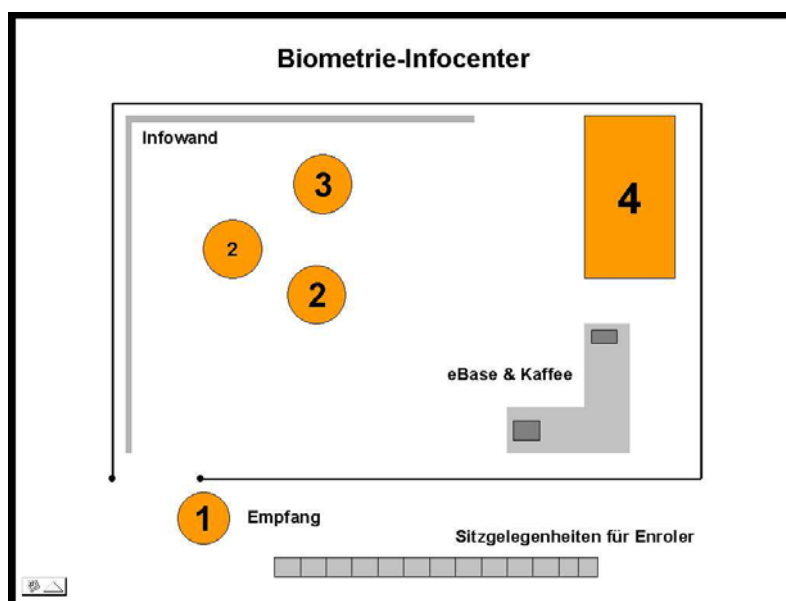


Abbildung 40: Enrolment-Center von Lufthansa im Biometrie-Infocenter

### 6.2.1.2 Enrolment-Personal

Das Enrolment-Personal wurde durch Lufthansa und Fraport bereitgestellt, wobei die Erstellung der Fotos des Gesichts durch Mitarbeiter des BKA erfolgte. Die Schulung und Einweisung des Enrolment-Personals übernahmen Mitarbeiter von secunet und des BKA. Am 11.03.2004 fand für das Enrolment-Personal der Lufthansa und Fraport jeweils eine Schulung im Umfang von sechs Stunden statt.

Die Schulung bestand aus einem theoretischen und einem praktischen Teil. Der theoretische Teil beinhaltete neben der Vorstellung der Projektpartner einen Überblick über das Projekt, eine

Einführung in das Thema Biometrie sowie wesentliche Aspekte des Enrolments. Bestandteile des praktischen Teils waren eine Einweisung an allen Systemen sowie die praktische Durchführung der verschiedenen Enrolment-Blöcke. Der gesamte Enrolment-Prozess wurde jeweils paarweise vom Enrolment-Personal eingeübt.

Die Teilnehmer der Schulung erhielten ein Handbuch, in dem alle wichtigen Informationen der Schulung zusammengefasst wurden. Das Handbuch enthielt darüber hinaus zu jedem biometrischen System eine Bedienungsanleitung zur Durchführung des Enrolments sowie die zugehörige Enrolment-Policy.

Schulung des Enrolment-Personals	
Theoretischer Teil	Praktischer Teil
<ul style="list-style-type: none"> <li>Vorstellung der Projektbeteiligten (BSI, BKA, Fraport, Lufthansa, secunet)</li> <li>Erläuterungen zur Frage „Was ist Biometrie?“</li> <li>Beschreibung des Testablaufs BioP II</li> <li>Beschreibung des Aufgaben des Enrolment-Personals</li> </ul>	<ul style="list-style-type: none"> <li>Empfang der Testpersonen und Ausgabe der RF-Token</li> <li>Erfassung der demographischen Daten</li> <li>Bedienung der biometrischen Systeme (Enrolment-Applikation)</li> </ul>

Tabelle 17: Schulung des Enrolment-Personals

### 6.2.1.3 Checkliste Enrolment

Das Enrolment-Personal von Fraport und Lufthansa wurde durch eine Checkliste in seiner Arbeit unterstützt. Dort wurde zunächst die UserID der Testperson ergänzt. Nach jedem Block wurde dann sukzessive die Checkliste mit dem jeweils erzielten Ergebnis ausgefüllt. Die Checkliste ermöglichte so auch eine Unterbrechung des Enrolments. Anhand der Checkliste konnte zu einem späteren Zeitpunkt zum Vervollständigen stets ermittelt werden, welcher Block des Enrolments noch fehlte.

Checkliste Enrolment UserID: 46990 \_\_\_\_\_

Biometrie-Infocenter	1. Versuch	2. Versuch	3. Versuch	
1. Ausgabe Token	<input type="checkbox"/> OK	---	---	
2. Erfassung statistischer Daten	<input type="checkbox"/> OK	---	---	
3. Bildaufnahme Finger				
Rechter Zeigefinger	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Ggf. Rechter Mittelfinger	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Linker Zeigefinger	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Ggf. Linker Mittelfinger	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
4. Bildaufnahme Gesicht	<input type="checkbox"/> OK	---	---	
<b>Teststation Tor 21</b>				
5. Enrolment „Cognitec“ (Gesicht)	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
6. Enrolment „SD Industries“ (Iris)				
Rechtes Auge	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Linkes Auge	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
7. Enrolment „BDr / NEC“ (Finger)				
Rechter Finger ( <b>gleicher Finger wie bei 3.</b> )	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Linker Finger ( <b>gleicher Finger wie bei 3.</b> )	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
8. Enrolment „Dermalog“ (Finger)				
Rechter Finger ( <b>gleicher Finger wie bei 3.</b> )	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK
Linker Finger ( <b>gleicher Finger wie bei 3.</b> )	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> OK	<input type="checkbox"/> Nicht OK

Abbildung 41: Checkliste Enrolment für das Enrolment-Personal Lufthansa

#### 6.2.1.4 Block 1 – Empfang und Ausgabe der RF-Karten

Für den Feldtest kamen RF-Karten des Typs Philips mifare 1k (mifare Standard Card IC MF1 IC S50) zum Einsatz. Die RF-Karten wurden mit gültigen UserIDs vorpersonalisiert, unternehmensspezifisch bedruckt (vergleiche Abbildung 42 und Abbildung 43) und in den Enrolment-Centern hinterlegt. Die RF-Karte wurde an eine Testperson ausgegeben, wenn diese ihre im Rahmen der Anmeldung zugeteilte UserID vorweisen konnte.



Abbildung 42: RF-Karte von Fraport

Abbildung 43: RF-Karte von Lufthansa

#### 6.2.1.5 Block 2 – Demographische Erfassung

Um selektive statistische Auswertungen des Feldtests zu ermöglichen, wurden ergänzend zu den biometrischen Merkmalen der Testteilnehmer auch demographische Daten erhoben (vergleiche Abbildung 44). Die Erhebung fand als eigener Block im Rahmen des Enrolments statt. Die Erhebung erfolgte anonym, da die erfassten Daten nur den UserIDs zugewiesen wurden. Die Erhebung erfolgte darüber hinaus freiwillig (jeweils Auswahlmöglichkeit „k. A.“).

User-ID		Geschlecht		Alter							Ethnische Herkunft								
		m	w	k. A.	unter 20	20-29	30-39	40-49	50-59	über 60	k. A.	euro-päisch	afrika-nisch	arabisch	asiatisch (indischer Subkont.)	asiatisch (Süd- u. Südostas.)	indianisch	andere	k. A.
46990	6070	5																	
46990	6078	3																	

Abbildung 44: Ausschnitt der Datei zur Erfassung der demographischen Daten (hier am Beispiel Fraport)

Der Umfang der erhobenen demographischen Daten wurde im Vorfeld zwischen allen Projektbeteiligten abgestimmt und von den betrieblichen Datenschutzbeauftragten und dem Betriebsrat von Lufthansa und Fraport genehmigt. Folgende demographische Daten wurden erhoben:

<b>Geschlecht</b>	<ul style="list-style-type: none"> <li>• männlich</li> <li>• weiblich</li> </ul>
<b>Alter</b>	<ul style="list-style-type: none"> <li>• unter 20</li> <li>• 20-29</li> <li>• 30-39</li> <li>• 40-49</li> <li>• 50-59</li> <li>• über 60</li> </ul>
<b>Ethnische Herkunft</b>	<ul style="list-style-type: none"> <li>• europäisch</li> <li>• afrikanisch</li> <li>• arabisch</li> <li>• asiatisch (indischer Subkontinent)</li> <li>• asiatisch (Süd- und Ostasien)</li> <li>• indianisch</li> <li>• andere</li> </ul>
<b>Berufsausbildung*</b>	<ul style="list-style-type: none"> <li>• keine</li> <li>• kaufmännische Ausbildung</li> <li>• handwerklich-technische Ausbildung</li> <li>• akademischer Grad (Natur- oder Ingenieurwissenschaften)</li> <li>• akademischer Grad (Humanwissenschaften, Jura, Management etc.)</li> <li>• andere</li> </ul>
<b>Berufliche Tätigkeit*</b>	<ul style="list-style-type: none"> <li>• operativ</li> <li>• administrativ</li> <li>• andere</li> </ul>
<b>Spezielle Merkmale*</b>	<ul style="list-style-type: none"> <li>• Kontaktlinsenträger (ja/nein) wenn ja: <ul style="list-style-type: none"> <li>○ harte Kontaktlinsen</li> <li>○ weiche Kontaktlinsen</li> <li>○ spezielle Kontaktlinsen</li> </ul> </li> <li>• Brillenträger (ja/nein)</li> <li>• Barträger (ja/nein) wenn ja: <ul style="list-style-type: none"> <li>○ Vollbart</li> <li>○ Teilbart</li> </ul> </li> </ul>

**Tabelle 18: Erhobene demographische Daten**

#### 6.2.1.6 Block 3 – Bildaufnahme der Fingerabdrücke (RefID3/5)

Mittels des optischen Sensors Sagem MorphoSmart™ MSO 100 wurde gemäß der Bedienungsanleitung (vergleiche Abbildung 45) das Bild je eines Fingers der rechten und linken Hand aufgenommen (Voreinstellung: Zeigefinger; alternativ Mittelfinger). Die Bildaufnahme musste einen voreingestellten Gütewert übersteigen, um akzeptiert zu werden. In der Checkliste Enrolment (vergleiche Abbildung 41) wurde festgehalten, von welchem Finger eine Bildaufnahme erstellt wurde.

\* Mehrfachnennungen möglich

In Block 5 (vergleiche Abschnitt 6.2.1.8) wurden dann dieselben Finger an den biometrischen Systemen enrolled.

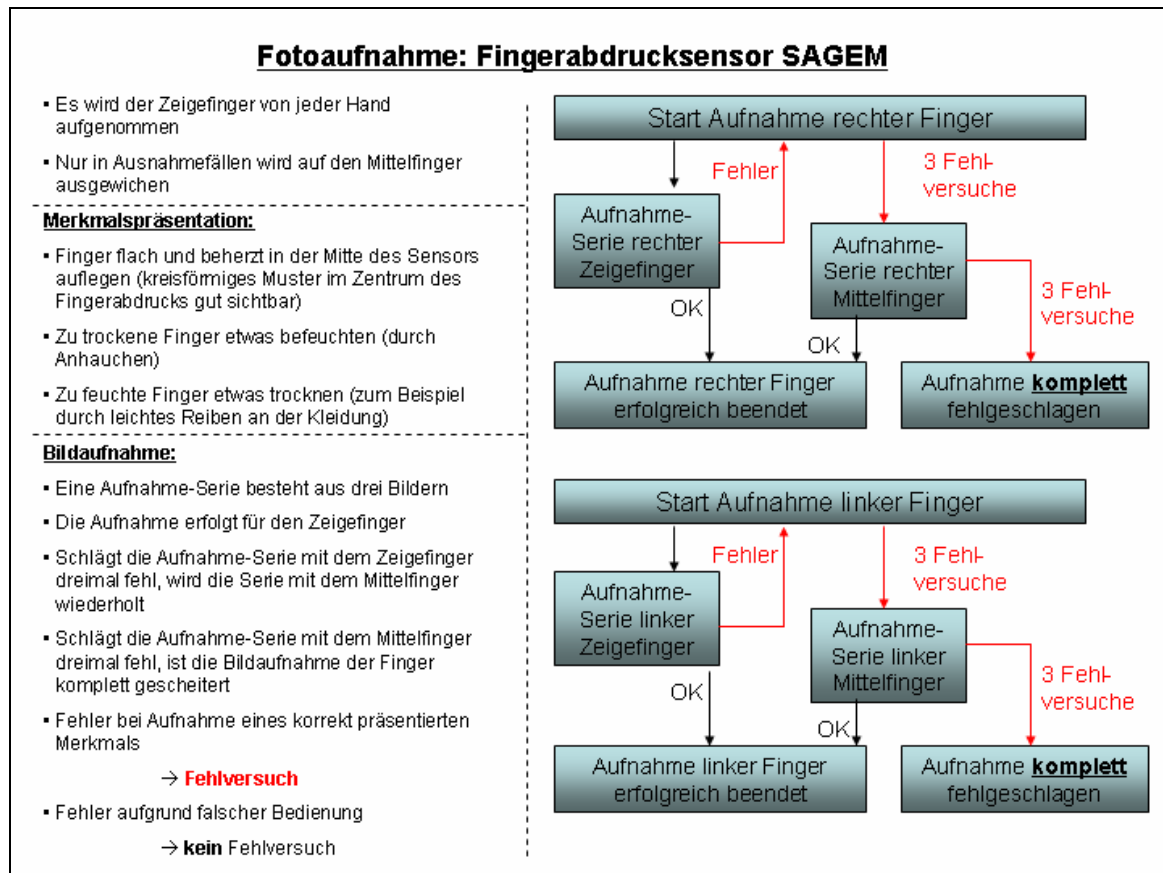


Abbildung 45: Kurzanleitung zur Bildaufnahme Fingerabdruck



**Abbildung 46: Bildaufnahme des rechten und linken Zeigefingers mit dem Sensor von Sagem**

#### 6.2.1.7 Block 4 – Erstellung der Passbilder (RefID1)

Mitarbeiter des BKA nahmen mit professioneller Ausrüstung ein Bild des Gesichts des Testteilnehmers auf. Dies erfolgt gemäß den Richtlinien der ICAO für die Erstellung von Passbildern zum Einsatz für biometrische Anwendungen (siehe Abschnitt 4.1.1).

#### 6.2.1.8 Block 5 – Erfassung der biometrischen Merkmale Gesicht (RefID2), Fingerabdruck (RefID4/6) und Iris (RefID8)

An dem Gesichtserkennungssystem von Cognitec wurde gemäß Enrolment-Policy (vergleiche Tabelle 19) das Gesicht der Testteilnehmer enrolt (RefID2). Anschließend erfolgte eine Testverifikation.

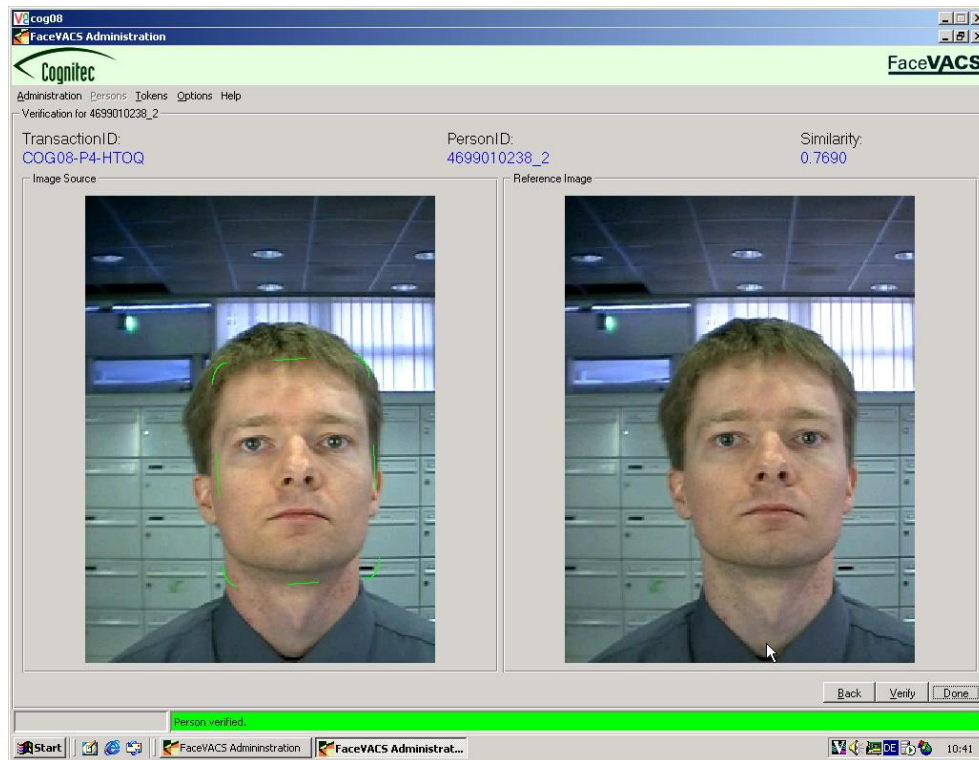


Abbildung 47: Enrolment am GE-System von Cognitec

<b>Qualitäts- bewertung</b>	<p>Nur ausreichend gute Aufnahmen des Merkmals dürfen als Referenzmuster gespeichert werden. Zur Bewertung der Aufnahmequalität müssen die vom System bereitgestellten Mechanismen genutzt werden (grüne umrahmte Bilder prüfen, ob das Gesicht der zu enrolenden Person richtig erfasst wurde).</p> <p>Die Qualitätsbewertung erfolgt also mittels einer Gütebewertung durch das System sowie zusätzlich durch eine Einschätzung des Enrolment-Administrators.</p>
<b>Template- Generierung</b>	Pro Testteilnehmer ein Template
<b>Merkmals- präsentation</b>	<ul style="list-style-type: none"> <li>• Abstand zum System durch Fußmarkierungen auf dem Boden</li> <li>• Neutraler Gesichtsausdruck (Augen geöffnet, Mund geschlossen)</li> <li>• Blick geradeaus</li> <li>• Leichtes Kreisen des Kopfes</li> <li>• Brillenträger behalten Brille auf</li> </ul>
<b>FTE</b>	<ul style="list-style-type: none"> <li>• Ein Enrolment-Versuch besteht aus der Aufnahme des Referenzmusters und – falls ein hinreichend gutes Referenzmuster erzeugt wurde – einem anschließenden Verifikationsversuch.</li> <li>• Ein Enrolment-Versuch gilt als nicht erfolgreich, wenn             <ul style="list-style-type: none"> <li>a) kein hinreichend gutes Referenzmuster gemäß Gütebewertung aufgenommen wurde oder</li> <li>b) bei dem folgenden Verifikationsversuch keine erfolgreiche Verifikation stattgefunden hat.</li> </ul> </li> <li>• Mindestens vier grün markierte Bilder sind für eine erfolgreiche Aufnahme des Merkmals notwendig.</li> <li>• Ein Fehler bei der Aufnahme des Merkmals (zum Beispiel das Minimum von 4 Bildern wird nicht erreicht) ergibt pro Fehler einen Enrolment-Fehlversuch. Fehler, die aufgrund falscher Bedienung entstehen, zählen hierbei nicht als Fehlversuch.</li> <li>• Eine Testverifikation ist nur dann als ordnungsgemäß durchgeführt zu betrachten, wenn das Gesicht korrekt präsentiert wurde.</li> <li>• Schlägt die Testverifikation trotz korrekt präsentiertem Merkmal fehl, bedeutet dies einen erneuten Enrolment-Fehlversuch.</li> <li>• Bei drei Fehlversuchen ist das Enrolment beendet.</li> <li>• Kann ein Teilnehmer nicht eingelernt werden, so muss dies im Enrolment-Protokoll (unter Angabe der Ursache) vermerkt werden.</li> </ul>

**Tabelle 19: Enrolment-Policy für die Erstellung der Gesichts-Templates (RefID2) am GE-System von Cognitec**

An dem Fingerabdruckererkennungssystem von Bundesdruckerei/NEC wurden gemäß Enrolment-Policy (vergleiche Tabelle 20) dieselben Finger enrolled, von denen in Block 3 die Bildaufnahmen erstellt wurden. Anschließend erfolgte eine Testverifikation je Finger.



Abbildung 48: Enrolment am FE-System von Bundesdruckerei/NEC

<b>Qualitäts- bewertung</b>	Nur ausreichend gute Aufnahmen des Merkmals dürfen als Referenzmuster gespeichert werden. Zur Bewertung der Aufnahmequalität müssen die vom System bereitgestellten Mechanismen genutzt werden (Gütebewertung erfolgt implizit durch das System – Prüfung Mindestanzahl von 20 Minuten).
<b>Template- Generierung</b>	Pro Testteilnehmer je ein Template des rechten und linken Zeigefingers (bzw. ersatzhalber des Mittelfingers)
<b>Merkmals- präsentation</b>	<ul style="list-style-type: none"> <li>• Das letzte Glied des Fingers ist bei jeder Aufnahme flach in der Mitte des Sensors aufzulegen, so dass das kreisförmige Muster im Zentrum eines Fingerabdrucks gut sichtbar ist.</li> <li>• Sehr trockene Finger, die nur schwache Linien auf dem Sensor erzeugen, müssen etwas befeuchtet werden (durch Anhauchen). Sehr feuchte Finger, die sehr starke Linien mit sichtbaren Schweißspuren auf dem Sensor erzeugen, müssen etwas getrocknet werden.</li> </ul>
<b>FTE</b>	<ul style="list-style-type: none"> <li>• Ein Enrolment-Versuch besteht aus der Aufnahme der Referenzmuster beider Finger und – falls ein hinreichend gutes Referenzmuster erzeugt wurde – einem anschließenden Verifikationsversuch pro Finger.</li> <li>• Ein Enrolment-Versuch gilt als nicht erfolgreich, wenn             <ul style="list-style-type: none"> <li>a) keine hinreichend guten Referenzmuster beider Finger gemäß Gütebewertung aufgenommen wurden oder</li> <li>b) bei dem folgenden Verifikationsversuch pro Finger keine erfolgreiche Verifikation statt gefunden hat.</li> </ul> </li> <li>• Das Enrolment wird mit dem rechten und linken Finger (Zeige- bzw. Mittelfinger) durchgeführt.</li> <li>• Ein Fehler bei der Aufnahme des Merkmals ergibt pro fehlerhafte Serie einen Enrolment-Fehlversuch. Fehler, die aufgrund falscher Bedienung entstehen, zählen hierbei nicht als Fehlversuch.</li> <li>• Wird ein einzelner Abdruck nicht innerhalb eines Zeitraumes von ca. 15 s als gültig erkannt, bedeutet dies ebenfalls einen Fehlversuch.</li> <li>• Eine Testverifikation ist nur dann als ordnungsgemäß durchgeführt zu betrachten, wenn der Finger korrekt präsentiert wurde.</li> <li>• Schlägt die Testverifikation trotz korrekt präsentiertem Merkmal fehl, bedeutet dies einen erneuten Enrolment-Fehlversuch.</li> <li>• Bei 3 Fehlversuchen ist das Enrolment beendet.</li> <li>• Kann ein Teilnehmer nicht eingelernt werden, so muss dies im Enrolment-Protokoll (unter Angabe der Ursache) vermerkt werden.</li> </ul>

**Tabelle 20: Enrolment-Policy für die Erstellung der Fingerabdruck-Templates (RefID4)  
am FE-System von Bundesdruckerei/NEC**

An dem Fingerabdruckererkennungssystem von Dermalog wurden gemäß der Enrolment-Policy (vergleiche Tabelle 21) dieselben Finger enrolled, von denen in Block 3 die Bildaufnahmen erstellt wurden (RefID3/5). Anschließend erfolgte eine Testverifikation je Finger.

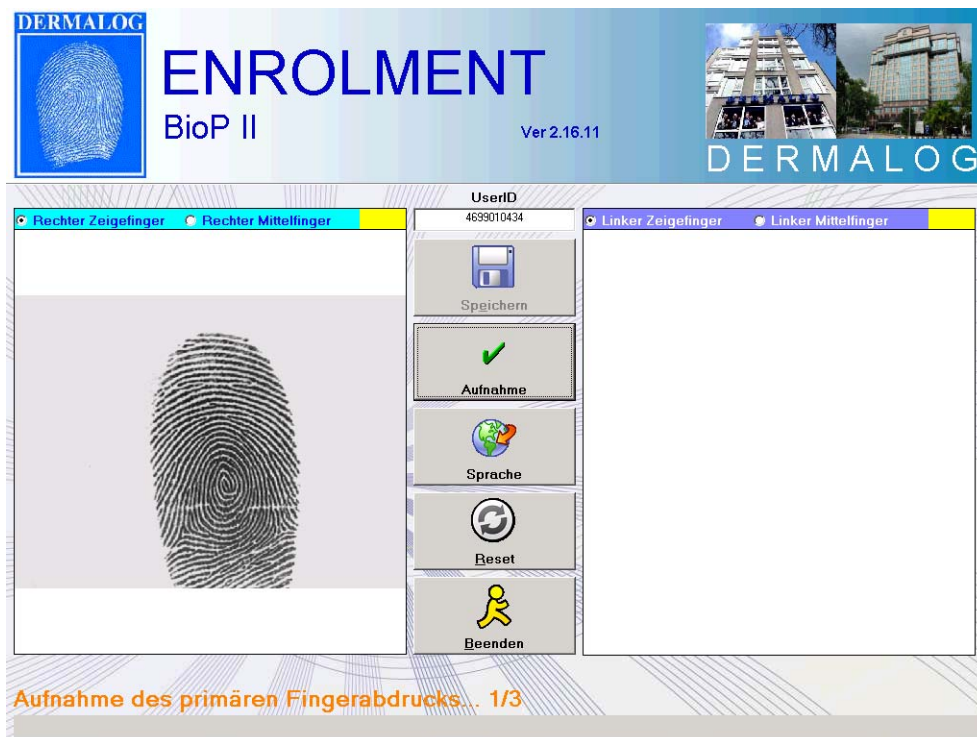


Abbildung 49: Enrolment am FE-System von Dermalog

<b>Qualitäts- bewertung</b>	Nur ausreichend gute Aufnahmen des Merkmals dürfen als Referenzmuster gespeichert werden. Zur Bewertung der Aufnahmequalität muss eine Gütebewertung durch Einschätzung des Enrolment-Administrators erfolgen (das kreisförmige Muster im Zentrum eines Fingerabdrucks muss gut sichtbar sein).
<b>Template- Generierung</b>	Pro Testteilnehmer je ein Template des rechten und linken Zeigefingers (bzw. ersatzhalber des Mittelfingers)
<b>Merkmals- präsentation</b>	<ul style="list-style-type: none"> <li>• Das letzte Glied des Fingers ist bei jeder Aufnahme flach in der Mitte des Sensors aufzulegen, so dass das kreisförmige Muster im Zentrum eines Fingerabdrucks gut sichtbar ist.</li> <li>• Sehr trockene Finger, die nur schwache Linien auf dem Sensor erzeugen, müssen etwas befeuchtet werden (durch Anhauchen). Sehr feuchte Finger, die sehr starke Linien mit sichtbaren Schweißspuren auf dem Sensor erzeugen, müssen etwas getrocknet werden.</li> </ul>
<b>FTE</b>	<ul style="list-style-type: none"> <li>• Ein Enrolment-Versuch besteht aus der Aufnahme der Referenzmuster beider Finger und – falls ein hinreichend gutes Referenzmuster erzeugt wurde – einem anschließenden Verifikationsversuch pro Finger.</li> <li>• Ein Enrolment-Versuch gilt als nicht erfolgreich, wenn             <ul style="list-style-type: none"> <li>a) keine hinreichend guten Referenzmuster beider Finger gemäß Gütebewertung aufgenommen wurden oder</li> <li>b) bei dem folgenden Verifikationsversuch pro Finger keine erfolgreiche Verifikation statt gefunden hat.</li> </ul> </li> <li>• Das Enrolment wird mit dem rechten und linken Finger (Zeige- bzw. Mittelfinger) durchgeführt.</li> <li>• Ein Fehler bei der Aufnahme des Merkmals ergibt pro fehlerhafte Serie einen Enrolment-Fehlversuch. Fehler, die aufgrund falscher Bedienung entstehen, zählen hierbei nicht als Fehlversuch.</li> <li>• Wird ein einzelner Abdruck nicht innerhalb eines Zeitraumes von ca. 15 s als gültig erkannt, bedeutet dies ebenfalls einen Fehlversuch.</li> <li>• Eine Testverifikation ist nur dann als ordnungsgemäß durchgeführt zu betrachten, wenn der Finger korrekt präsentiert wurde.</li> <li>• Schlägt die Testverifikation trotz korrekt präsentiertem Merkmal fehl, bedeutet dies einen erneuten Enrolment-Fehlversuch</li> <li>• Bei 3 Fehlversuchen ist das Enrolment beendet.</li> <li>• Kann ein Teilnehmer nicht eingelernt werden, so muss dies im Enrolment-Protokoll (unter Angabe der Ursache) vermerkt werden.</li> </ul>

**Tabelle 21: Enrolment-Policy für die Erstellung der Fingerabdruck-Templates (RefID6) am FE-System von Dermalog**

Am Iriserkennungssystem von SDI wurden gemäß der Enrolment-Policy (siehe Tabelle 22) die rechte und linke Iris der Testteilnehmer enrolt. Anschließend erfolgte je eine Testverifikation pro Iris.



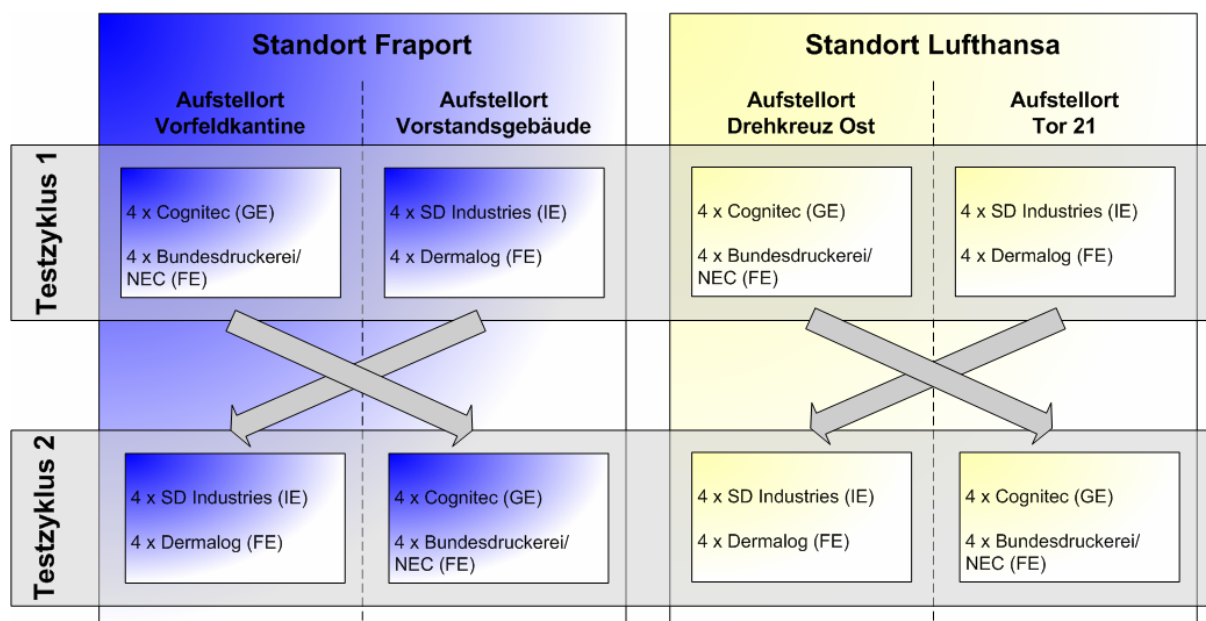
Abbildung 50: Enrolment am IE-System von SD Industries

<b>Qualitäts- bewertung</b>	Nur ausreichend gute Aufnahmen des Merkmals dürfen als Referenzmuster gespeichert werden. Zur Bewertung der Aufnahmequalität müssen die vom System bereitgestellten Mechanismen genutzt werden (Gütebewertung erfolgt implizit durch das System).
<b>Template- Generierung</b>	Pro Testteilnehmer je ein Template des rechten und linken Auges
<b>Merkmals- präsentation</b>	<ul style="list-style-type: none"> <li>• Positionierung ca. 20 cm vom System entfernt</li> <li>• Zu enrolendes Auge weit öffnen, das andere Auge zuhalten</li> <li>• Mit dem zu enrolenden Auge in den Spiegel schauen</li> <li>• Brille absetzen, Kontaktlinsen können anbehalten werden; bei Verwendung von Speziallinsen (farbig etc.) muss dies vermerkt werden</li> </ul>
<b>FTE</b>	<ul style="list-style-type: none"> <li>• Ein Enrolment-Versuch besteht aus der Aufnahme der Referenzmuster beider Augen und – falls ein hinreichend gutes Referenzmuster erzeugt wurde – einem anschließenden Verifikationsversuch pro Auge</li> <li>• Ein Enrolment-Versuch gilt als nicht erfolgreich, wenn             <ul style="list-style-type: none"> <li>a) keine hinreichend guten Referenzmuster beider Augen gemäß Gütebewertung aufgenommen wurden oder</li> <li>b) bei dem folgenden Verifikationsversuch pro Auge keine erfolgreiche Verifikation statt gefunden hat.</li> </ul> </li> <li>• Das Enrolment wird mit dem rechten und linken Auge durchgeführt.</li> <li>• Ein Fehler bei der Aufnahme des Merkmals (zum Beispiel wenn die Fehlermeldung „Bad Image Quality“ erscheint) ergibt pro Fehler einen Enrolment-Fehlversuch. Fehler, die aufgrund falscher Bedienung entstehen, zählen hierbei nicht als Fehlversuch.</li> <li>• Eine Testverifikation ist nur dann als ordnungsgemäß durchgeführt zu betrachten, wenn das Auge korrekt präsentiert wurde.</li> <li>• Schlägt die Testverifikation trotz korrekt präsentiertem Merkmal fehl, bedeutet dies einen erneuten Enrolment-Fehlversuch.</li> <li>• Bei 3 Fehlversuchen ist das Enrolment beendet.</li> <li>• Kann ein Teilnehmer nicht eingelernt werden, so muss dies im Enrolment-Protokoll (unter Angabe der Ursache) vermerkt werden.</li> </ul>

**Tabelle 22: Enrolment-Policy für die Erstellung der Iris-Templates (RefID8) am IE-System von SD Industries**

### 6.2.2 Feldtest

Der Feldtest gliederte sich in zwei Testzyklen zu je ca. acht Wochen. Diese Aufteilung war erforderlich, damit durch einen Systemtausch zwischen den Aufstellorten eines Standorts eine gleichmäßige Bedienung aller vier biometrischen Systeme durch alle Testteilnehmer erfolgen konnte (vergleiche Abbildung 51).



**Abbildung 51: Aufteilung der biometrischen Systeme nach Testzyklen**

#### 6.2.2.1 Änderungen an Konfiguration und Umgebung während des Feldtests

Durch den Wechsel wurden die Umgebungsbedingungen in den Einsatzorten bis auf den Austausch der Systeme selbst nicht variiert. Des Weiteren wurden keine Konfigurationsänderungen an den biometrischen Systemen vorgenommen.

Eine Ausnahme stellen in diesem Zusammenhang die GE-Systeme von Cognitec dar. Anhand der Ergebnisse der Gesichtserkennung in Testzyklus 1 stellte sich heraus, dass die Rückweisungsrate für ICAO-Bilder weit höher war als entsprechend der Erfahrungen mit Cognitec aus BioP I erwartet worden war. Eine ausführliche Analyse der Ergebnisse mit Einbeziehung des Herstellers zeigte, dass die Umgebungsbedingungen an den Einsatzorten zu einer schlechten Erkennungsleistung auf Basis der ICAO-Bilder führten, obwohl die diesbezüglichen Herstellerempfehlungen berücksichtigt wurden (zum Beispiel keine direkte Sonneneinstrahlung, Ersetzung der ursprünglichen Beleuchtung durch indirekt abstrahlende Deckenfluter, Zusatzbeleuchtung an den Systemen). Als wesentliche Ursachen für die hohen Rückweisungsrate stellten sich eine nicht ausreichend homogene Ausleuchtung des Gesichts sowie Blendstellen auf Brillen durch die am Cognitec-System angebrachte Beleuchtung heraus. Aus diesem Grund wurde die diffuse Grundhelligkeit im Raum erhöht. Hierzu wurden auf den GE-Systemen zusätzliche Lampen bzw. Strahler angebracht, so dass die Gesichter der Benutzer stärker indirekt beleuchtet werden (Lichtreflexion über die Raumdecke und über die Rück- und Seitenwände). Zusätzlich wurden am Aufstellort Fraport Vorstandsgebäude die ursprünglich vorhandenen grauen Gitterdeckenplatten, die kaum reflektierten, durch helle, diffus reflektierende Deckenplatten ersetzt. Außerdem wurde die direkte Beleuchtung reduziert, um Reflexionen in Brillengläsern zu vermeiden. Dazu wurde die Intensität der an den GE-Systemen angebrachten direkten Beleuchtung (vertikal angebrachte Stablampen) reduziert.

Neben den genannten Änderungen der Umgebung für die Cognitec-Systeme wurden die Toleranzschwelle und die Masterreferenz geändert (siehe Abschnitt 6.1.6.1). Die Reduzierung der Toleranzschwelle erwies sich auf Basis der Analyse der Ergebnisse aus Testzyklus 1 als sinnvoll. Mit dem dort eingestellten Niveau liefen die GE-Systeme auf einem sehr hohen Sicherheitsniveau. Durch die Umstellung der Masterreferenz auf ICAO-Bild lieferte das System für diesen Vergleich optimierte Live-Bilder.

Durch die Konfigurations- und Umgebungsänderungen für die GE-Systeme ist keine Vergleichbarkeit der Ergebnisse aus Testzyklus 1 und 2 gegeben. Aus diesem Grund wird in den BioP II Auswertungen für das ICAO-Bild jeweils nur Testzyklus 2 berücksichtigt. Auswertungen auf Basis des Templates erfolgen dagegen für Testzyklus 1, da hier die Testumgebung der Enrolment-Umgebung entspricht und für Testzyklus 2 kein erneutes Enrolment durchgeführt werden konnte.

Die an den Dermalog-Systemen vorgenommene Anpassung der Sensorempfindlichkeit (siehe Abschnitt 6.1.6.3) diente lediglich dazu, die Bedienerfreundlichkeit des Systems zu verbessern.

#### 6.2.2.2 Benutzertraining während des Feldtests

Erfahrungen haben gezeigt, dass in den ersten Tagen eines Feldtests aufgrund der Ungeübtheit der Teilnehmer signifikant schlechtere Ergebnisse auftreten. Außerdem ist oft noch eine Feinabstimmung der biometrischen Systeme erforderlich. Daher wurde jedem Testzyklus eine mindestens einwöchige Startphase (TeachIn) vorgeschaltet.<sup>26</sup> Der Testablauf in dieser Einführungsphase ist identisch mit dem in der regulären Testphase. Zusätzlich wurden für die Startphase folgende Festlegungen getroffen:

- **Benutzerführung:** Beim TeachIn wurde eine intensive Betreuung der Benutzer sichergestellt. Es wurde gewährleistet, dass bei der Bedienung der biometrischen Systeme durch die Testteilnehmer Personen anwesend sind, die Hinweise und Hilfestellungen gaben. Das TeachIn wurde im Wesentlichen von Mitarbeitern der secunet und des BKA durchgeführt.
- **Ergebnisse:** In der Einführungsphase wurden die anfallenden Ergebnisse genauso erfasst wie im regulären Test. Die Ergebnisse gingen jedoch nicht in die Feldtestauswertung und somit in die Bewertung der Systeme ein. Die Ergebnisse wurden jedoch dahingehend untersucht, ob ggf. Probleme auftraten. Diese wurden dann soweit möglich vor dem Start der regulären Testphase behoben.
- **Anpassung:** Die Systeme wurden auf möglicherweise erforderliche Feinabstimmungen geprüft und diese soweit notwendig vorgenommen.

#### 6.2.2.3 Re-Enrolment

Insbesondere während der TeachIn-Phase und während der ersten Tage des Feldtests wurde darauf geachtet, ob einzelne Personen häufiger als andere zurückgewiesen wurden. Sofern kein fehlerhaftes Verhalten seitens der Teilnehmer als Ursache hierfür festgestellt werden konnte, wurde ein Re-Enrolment durchgeführt. Dies verteilte sich wie folgt:

- Cognitec 52 Re-Enrolments
- Bundesdruckerei/NEC 23 Re-Enrolments
- Dermalog 104 Re-Enrolments
- SD Industries 13 Re-Enrolments

Die vergleichsweise hohen Zahlen von Re-Enrolments bei Cognitec und Dermalog sind darauf zurückzuführen, dass aufgrund von Systemfehlern während des Enrolments teilweise keine korrekte Protokollierung erfolgte und somit das Template des betroffenen Testteilnehmers nicht in der Referenzdatenbank vorlag.

---

<sup>26</sup> Sieben Werktagen in Zyklus 1, fünf Werktagen in Zyklus 2

## 7 Feldtestauswertung

### 7.1 Auswertungskonzept

Nachfolgend wird das Auswertungskonzept von BioP II dargestellt, das die verschiedenen Projektziele berücksichtigt. Die Auswertung orientiert sich bezüglich Begriffen und Methoden an [ISO\_BPTR], sofern dort entsprechende Themen behandelt werden. Abweichungen erfolgen nur in begründeten Fällen und werden gekennzeichnet. Neben den Untersuchungen der technischen Leistungsfähigkeit gemäß [ISO\_BPTR] beinhaltet BioP II Untersuchungen hinsichtlich Zuverlässigkeit und sicherheitstechnischer Eigenschaften der biometrischen Systeme sowie menschlicher Faktoren wie Bedienbarkeit und Benutzerakzeptanz.

#### 7.1.1 Vergleichstypen

Die Zielsetzung von BioP II beinhaltet verschiedene Vergleichstypen (siehe Abbildung 52):

- Vergleich der biometrischen Verfahren Gesichts-, Fingerabdruck- und Iriserkennung anhand der sie repräsentierenden Systeme
- Vergleich unterschiedlicher Referenzbasen (Bild gemäß ICAO-Richtlinien vs. system-spezifisches Template) für das jeweilige System

System	Cognitec	Bundesdr./ NEC	Dermalog	SD Industr.
Referenz				
ICAO	←————— 1 —————→			
TPL	↕ 2 ↕			

Abbildung 52: Vergleichstypen in BioP II

Gemäß [ISO\_BPTR] ist der Systemvergleich und somit der Feldtest, wie er sich für die Teilnehmer darstellt, eine „Szenario-Evaluation“. Detailanalysen im Rahmen der weiterführenden Untersuchungen sind als „Technische Evaluation“ gemäß [ISO\_BPTR] zu klassifizieren.

#### 7.1.2 Bewertung der Erkennungsleistungen

Wesentliches Kriterium zur Bewertung von biometrischen Systemen ist deren Erkennungsleistung. Diese ergibt sich aus den Wahrscheinlichkeiten, dass eine berechtigte Person vom System zurückgewiesen wird und dass eine unberechtigte Person vom System akzeptiert wird. Diese Wahrscheinlichkeiten werden als False Reject Rate (FRR) und False Accept Rate (FAR) bezeichnet. Die Werte für FRR und FAR sind nicht theoretisch herleitbar, sondern müssen immer statistisch auf Basis aufwendiger Tests ermittelt werden.

Da die Werte für FRR und FAR korrelieren, müssen für die Bewertung der Erkennungsleistung eines Systems immer beide Werte angegeben werden. Singuläre Angaben von FAR oder FRR sind nicht sinnvoll (da dann jeweils gute Arbeitspunkte herausgesucht werden können). So ist beispielsweise die

Angabe einer niedrigen FRR nicht aussagekräftig, wenn der zugehörige Wert der FAR (der dann in der Regel hoch ist) nicht bekannt ist.

Per Definition berechnet sich eine FRR immer aus Verifikationen berechtigter Personen, die FAR dagegen aus Verifikationen Unberechtigter. Dies bedeutet für BioP II, dass zur Angabe von Erkennungsleistungen beide Vorgänge in hoher Anzahl durchzuführen sind, um statistisch signifikante Aussagen treffen zu können.

Zur Bestimmung der FRR finden die Verifikationen der Feldtestteilnehmer über den Feldtestzeitraum Verwendung. Die Ermittlung der FAR erfolgt dagegen auf Basis von im Feldtest gewonnenen Live-Bildern, die nach dem Feldtest im Batch-Modus zur Verifikation gegenüber den Referenzen anderer Personen genutzt werden. Damit wird eine große Anzahl von Verifikationen Unberechtigter simuliert.

Für die Darstellung der Erkennungsleistung von biometrischen Systemen gibt es mehrere Verfahren. Für BioP II finden drei verschiedene Methoden Verwendung, die nachfolgend erläutert werden.

Eine einfache und übersichtliche Methode ist die **Darstellung der relativen Häufigkeiten von Matchscores** (Genuine-Impostor-Frequency-Diagramm). Ein Matchscore ist der erzielte Trefferwert beim Vergleich des aktuell aufgenommenen Merkmals der zu authentifizierenden Person mit dem zugehörigen gespeicherten Referenzmerkmal. Berechtigte Personen erzielen typischerweise hohe Matchscores, Unberechtigte in der Regel niedrige. Eine Ausnahme bildet die Iriserkennung. Hier sind die Matchscores umso niedriger je höher der Übereinstimmungsgrad ist.

Bei dieser Form der Darstellung werden auf der Abszisse alle im Test angefallenen Werte von Matchscores aufgetragen. Die Ordinate enthält die jeweils zugehörigen relativen Häufigkeiten des Auftretens (also die absolute Anzahl des Auftretens eines Matchscore-Wertes normiert bezüglich der Gesamtzahl aller Matchscores). Diese Werte werden jeweils für Berechtigte und Unberechtigte aufgetragen. Im Idealfall zeigen beide Verteilungskurven keine Überschneidung.

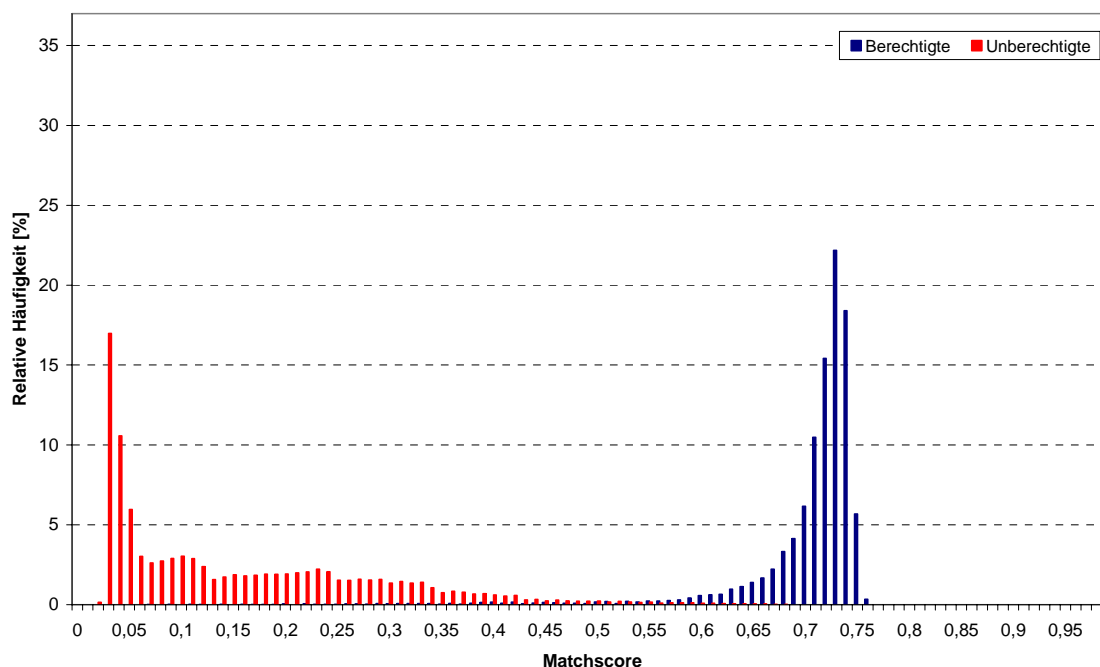
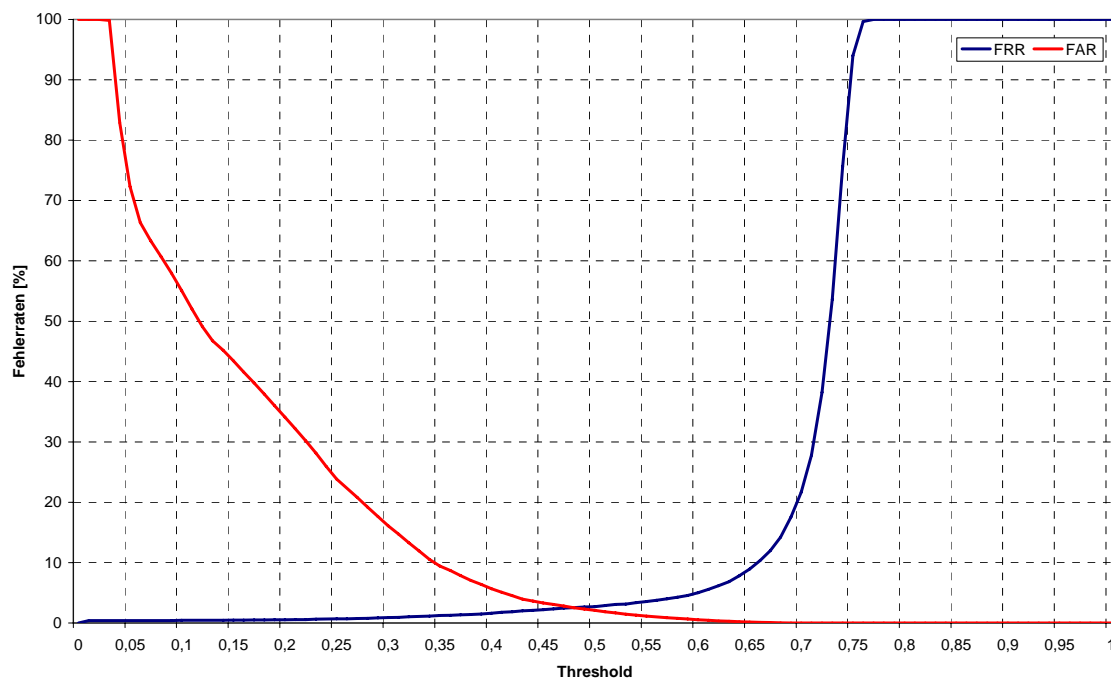


Abbildung 53: Beispiel eines Genuine-Impostor-Frequency-Diagramm

Aus diesem Histogramm lassen sich recht einfach FAR und FRR ableiten. Legt man einen Matchscore als Toleranzschwelle (Threshold) zur Unterscheidung von Berechtigten und Unberechtigten fest, dann ergibt sich die FAR aus der Anzahl der oberhalb (für Iriserkennung unterhalb) dieser Schwelle liegenden Matchscores von Unberechtigten im Verhältnis zur Gesamtanzahl der Versuche bzw. angefallenen Matchscores. Umgekehrt ergibt sich die FRR aus der unterhalb (für Iriserkennung oberhalb) der Schwelle liegenden Anzahl von Matchscores von Berechtigten im Verhältnis zur zugehörigen Gesamtanzahl. Somit können sukzessive aus diesen Verteilungskurven FAR-FRR-Kurven ermittelt werden. Diese stellen die Fehlerraten in Abhängigkeit des Schwellwertes dar.

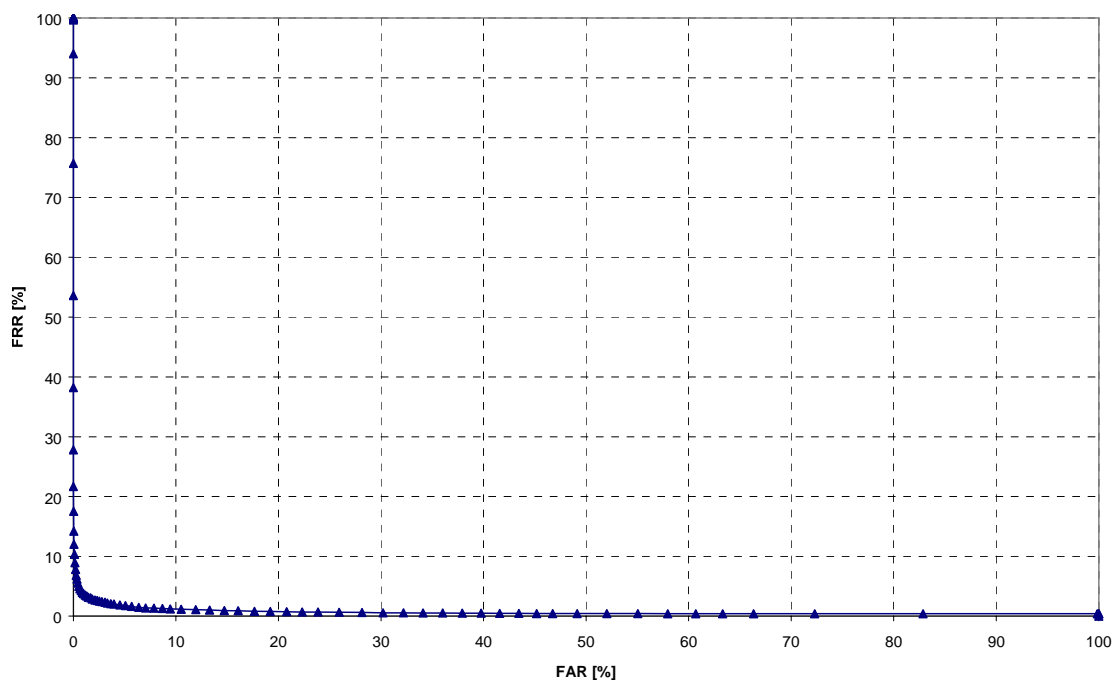


**Abbildung 54: Beispiel eines FAR-FRR-Diagramm**

**FAR-FRR-Diagramme** sind die verbreitetste Darstellung der Erkennungsleistungen eines biometrischen Systems. Da sie leicht verständlich sind, werden sie auch für BiOP II verwendet. Diese Darstellung ist insbesondere für die Festlegung eines Thresholds eines Systems für einen bestimmten Einsatzzweck geeignet. Absolute Aussagen über die tatsächliche Leistungsfähigkeit des Systems und insbesondere der Vergleich verschiedener biometrischer Systeme sind mit diesem Diagrammtyp nur beschränkt möglich. Das liegt vor allem daran, dass auftretende Matchscores immer algorithmenspezifisch realisiert sind. Damit sind die Matchscores und entsprechend die resultierenden Thresholds für unterschiedliche Systeme nicht vergleichbar. Für die Verteilung der Matchscores sind beliebige Skalierungen und Transformationen realisierbar, die das Aussehen von FAR-FRR-Kurven beeinflussen. Häufig werden beispielsweise einzelne Arbeitsbereiche der Kurve gedehnt, um das System robuster gegen Schwellwertänderungen erscheinen zu lassen. Diese Methoden beeinflussen jedoch nicht das Verhältnis der Werte FAR und FRR zueinander. Dies motiviert die direkte Darstellung der FRR in Abhängigkeit der FAR. Damit wird der Threshold als Parameter eliminiert und es wird eine von Schwellwertskalierungen unabhängige Darstellung erreicht, die eine echte Vergleichbarkeit unterschiedlicher biometrischer Systeme oder auch Systemkonfigurationen ermöglicht.

Diese Darstellung kann mittels einer **DET-Kurve** (Detection Error Trade-off) erfolgen. Dabei wird die FRR als Funktion der FAR dargestellt. Die ideale DET-Kurve nimmt nur Werte auf den

Koordinatenachsen an ( $FAR \neq 0 \Rightarrow FRR = 0$  und umgekehrt). Der oberste Punkt ist für alle Systeme durch  $FAR = 0\%$  und  $FRR = 100\%$  gegeben (DET-Kurven können per Definition nicht ansteigen). Generell gilt, je enger sich die Kurve eines Systems an die Achsen anschmiegt, umso besser ist die Erkennungsleistung. Die Equal Error Rate (EER), bei der  $FAR = FRR$  gilt, leitet sich aus dem Schnittpunkt der DET-Kurve mit der Diagonalen des Koordinatensystems ab.



**Abbildung 55: Beispiel eines DET-Diagramm**

In BioP II werden DET-Kurven ermittelt, indem für den gesamten Bereich möglicher Thresholds die jeweils zugehörige FAR und FRR berechnet wird. Damit wird die vollständige Leistungsfähigkeit dargestellt. Ein Vergleich verschiedener Systeme ist damit einfach möglich, indem verschiedene Kurven innerhalb eines Diagramms dargestellt werden. Da sich die resultierenden Kurven nicht notwendig über den ganzen Arbeitsbereich so anordnen, so dass ein einfaches Ranking möglich wäre, kann die Erkennungsleistung nur anhand einzelner Arbeitspunkte bewertet werden.

Eine Alternative zur DET-Kurve ist die Darstellung mittels ROC-Kurve. Hier wird auf der Ordinate ( $1 - FRR$ ) statt FRR aufgetragen.

Des Weiteren werden alternativ zu FRR bzw. FAR auch FNMR (False Non Match Rate) und FMR (False Match Rate) verwendet. Dies ist insbesondere beim Vergleich isolierter biometrischer Algorithmen auf Basis innerhalb von Datenbanken vorliegender Eingabebilder üblich. FRR und FNMR bzw. FAR und FMR unterscheiden sich darin, dass FRR und FAR Fehler bei der Aufnahme des Live-Bildes bzw. der Template-Generierung beinhalten. Dieser Fehler wird in der FTA (Failure-To-Acquire Rate) zusammengefasst. Die FTA ist der Anteil von Verifikationsversuchen, bei denen das System nicht in der Lage ist, ein Bild ausreichender Qualität zum Matching zur Verfügung zu stellen. Mögliche Ursachen sind, dass das Merkmal nicht präsentiert werden kann (zum Beispiel wegen Verletzung), die Merkmalsextraktion scheitert oder eine Qualitätsschwelle (zum Beispiel ausreichende Anzahl Minuten) nicht erreicht wird. Die FRR bzw. FAR berechnen sich aus FNMR bzw. FMR gemäß nachfolgenden Gleichungen.

$$FRR = FTA + FNMR(1 - FTA)$$

$$FAR = FMR(1 - FTA)$$

Gemäß [ISO\_BPTR] werden FNMR und FMR als Matching Error und FRR und FAR als Decision Error zusammengefasst.

Unter den Ausgangsbedingungen, dass in BioP II Komplettsysteme in einem simulierten Zielszenario bewertet werden (Szenario-Evaluation) und die Tests auf Basis echter Versuche durch Teilnehmer erfolgen, werden in BioP II ausschließlich Decision Errors ermittelt. Dies entspricht der Empfehlung von [ISO\_BPTR], entsprechende Systemvergleiche auf Basis von Decision Error DETs durchzuführen.

In BioP II werden keine kumulierten FRRs<sup>27</sup> bestimmt (zum Beispiel die Rückweisungsrate nach der zweiten Betätigung), sondern nur die Rückweisungsrate bei der ersten Betätigung. Die Durchführung mehrerer Erkennungsversuche ist bereits durch die Funktionsweise der beteiligten Systeme bedingt. Bei Nichterkennung werden automatisch neue Versuche bis zur Erreichung eines Timeout angestoßen (siehe Abschnitt 6.1.6).

#### 7.1.2.1 Systemvergleich

Die Erkennungsleistung eines biometrischen Systems muss immer in der Kombination FRR und zugehöriger FAR angegeben werden. Um die beteiligten Systeme vergleichen zu können, werden hierzu Arbeitspunkte festgelegt. Diese Festlegung kann sich an fixen Werten für die FAR oder für die FRR orientieren. Für BioP II wurden folgende Punkte betrachtet:

- FAR = 0,001%
- FAR = 0,01%
- FAR = 0,1%
- FAR = 1%

Bei einem Vergleich muss jeweils ein FAR-Wert ausgewählt und für das betreffende System die korrespondierende FRR berechnet bzw. abgelesen werden (vergleiche Abbildung 56). Das Ergebnis kann dann zum Vergleich herangezogen werden.

---

<sup>27</sup> Kumuliert im Sinne von Zusammenfassung mehrerer Betätigungen zu einem Versuch

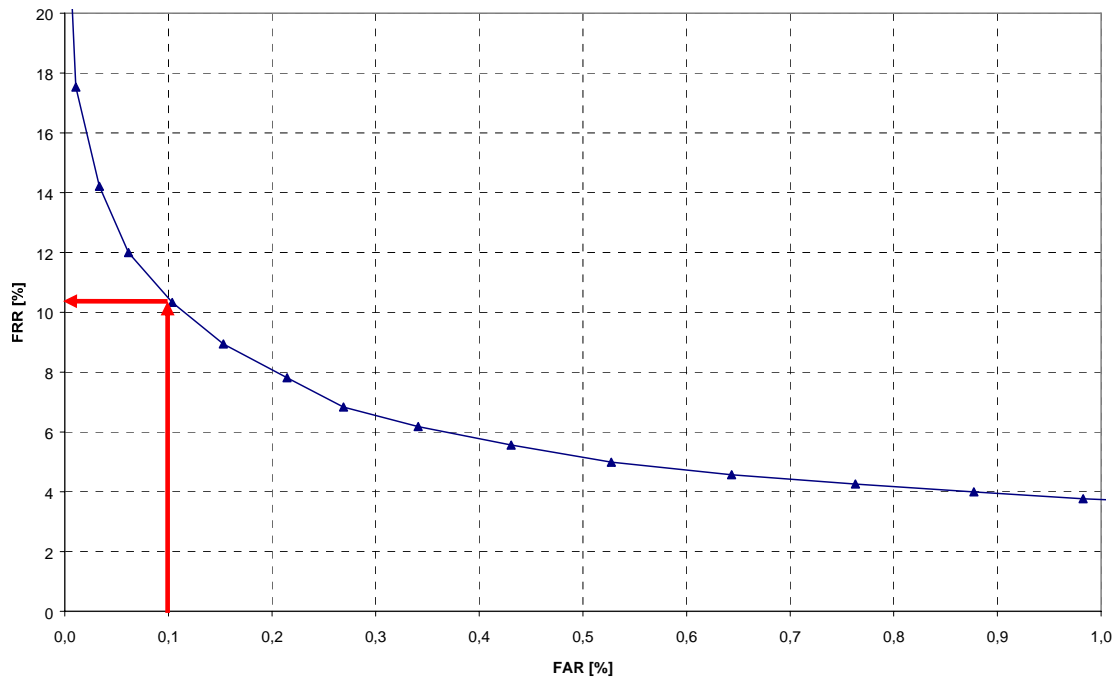


Abbildung 56: Arbeitspunkt FAR = 0,1%

Der Schwerpunkt der Untersuchungen in BioP II liegt auf dem Arbeitspunkt FAR = 0,1%.

#### 7.1.2.2 Auswertungstypen

Die biometrischen Systeme wurden während des Feldtests unterschiedlich häufig durch die Testteilnehmer bedient. Um entsprechende Auswirkungen auf die Erkennungsleistung zu berücksichtigen, wurden folgende Auswertungsmethoden angewandt.

- Für die Ermittlung der **Gesamterkennungsleistung** fanden alle Verifikationen der Feldtestteilnehmer über den Feldtestzeitraum Verwendung.
- Da durch unterschiedliche Bedienungshäufigkeiten der Teilnehmer einzelne Personen die Gesamterkennungsleistung signifikant beeinflussen können, wurden in BioP II zusätzlich **nutzerklassenspezifische Erkennungsleistungen** ermittelt. Hierbei gehen jeweils alle Verifikationen von Teilnehmern in die FRR-Berechnung ein, die einen bestimmten Bereich von Bedienungshäufigkeiten erreicht haben (vergleiche Definition Nutzerklassen in Abschnitt 6.1.1.2).
- Des Weiteren wurden auch **nutzerspezifische Erkennungsleistungen** – die so genannte **Einzelbenutzerstatistik** – ermittelt. Die Einzelbenutzerstatistik stellt spezifisch für jeden einzelnen Testteilnehmer die FRR dar. Idealerweise sollten die FRRs der verschiedenen Teilnehmer alle möglichst niedrig sein. Hohe FRRs einzelner Personen deuten darauf hin, dass entweder die Erfassung des Merkmals dieser Person ein Problem bereitet oder das Merkmal der Person vom Algorithmus bzgl. der betrachteten Referenzbasis nicht hinreichend gut verwertet werden kann.
- Um eine von der Betätigungshäufigkeit der Testteilnehmer unabhängige Kennzahl für die Gesamterkennungsleistung zu erhalten, ist eine **Normierung** der individuellen Erkennungsleistungen erforderlich. Eine Alternative hierfür ist die Bestimmung der FRR auf Basis unabhängiger Ereignisse. Das würde bedeuten, dass für jeden Teilnehmer genau ein zufällig

ausgewähltes Verifikationsergebnis in die Berechnung der FRR eingeht. Für BioP II findet eine andere Alternative Anwendung. Hier wird eine **gemittelte Einzelbenutzerstatistik** bestimmt, das heißt jeder Teilnehmer geht mit seiner spezifischen FRR gleich stark in das Gesamtergebnis ein.

Für alle vier beschriebenen Auswertungstypen wird jeweils für feste FAR-Arbeitspunkte die zugehörige FRR ermittelt.

## 7.2 Testergebnisse

### 7.2.1 Definition der Basisdatenmengen

#### 7.2.1.1 Verifikationsergebnisse zur FRR-Ermittlung

In BioP II wurden sämtliche Betätigungen in der Datenbank protokolliert, sofern die UserID der betreffenden Person bekannt war. Es ist daher eine Einschränkung der Gesamtmenge von Betätigungen auf Teilmengen erforderlich, die nur die für die Auswertung relevanten Betätigungen beinhalten.

Mögliche Kriterien zur Einschränkung der Gesamtmenge von Verifikationsdaten auf die auszuwertenden Teilmengen sind im Folgenden aufgezählt:

1. Betätigungen außerhalb des definierten Feldtestzeitraums
2. Betätigungen von Projektmitarbeitern
3. Durch signifikant variierte Testbedingungen beeinflusste Ergebnisse
4. Durchführung einer Folge von Betätigungen an einem System durch eine Person (bewirkt nicht zu vernachlässigende statistische Abhängigkeiten nach der ersten Betätigung)
5. Nicht komplett vorliegende Referenzdaten zu einer Person (zum Beispiel bei Iris oder Finger nur ein Merkmal enrolbar)
6. Person hat nur sehr wenige Betätigungen durchgeführt (Bedienfehler dominieren gegenüber Erkennungsfehlern)
7. Person hat überdurchschnittlich viele Betätigungen durchgeführt (ohne Normierung ist der Einfluss dieser Person auf das Gesamtergebnis zu stark)
8. Vertauschung von RF-Token durch Feldtestteilnehmer (Spieltrieb, Überwindungsversuch)
9. Nicht kooperatives Verhalten der Testteilnehmer (bewusst variiertes Aussehen durch Grimassen, Sonnenbrille etc., Verwendung eines nicht eingelernten Fingers etc.)
10. Fehlfunktion des Systems

Für die in den folgenden Kapiteln dargestellten Ergebnisse erfolgt lediglich ein **Ausschluss von Betätigungen gemäß den Kriterien 1 bis 5**. Die Kriterien 6 und 7 führen nicht zu einem Ausschluss der Betätigung, sondern zu einer Betrachtung in verschiedenen Nutzerklassen (siehe Abschnitt 6.1.1.2). Die Kriterien 8 bis 10 wurden für eine Abschätzung des aus ihnen resultierenden Fehlers stichprobenartig für einzelne Feldtesttage komplett analysiert.

Bei den Fingerabdruck- und den Iriserkennungssystemen wurden für jede Referenzbasis pro Betätigung zwei Verifikationsergebnisse protokolliert: zum einen das Ergebnis des 1:1-Vergleichs mit dem primären Referenzmerkmal und zum anderen das Ergebnis des 1:2-Vergleichs mit beiden Referenzmerkmalen (primär und sekundär). Da eine nicht unerhebliche Anzahl von Teilnehmern häufig das sekundäre Merkmal bei einer Betätigung präsentierte (für BDr/NEC bei ca. 7,5% aller

Betätigungen, für Dermalog bei ca. 5% aller Betätigungen, für SDI bei ca. 23% aller Betätigungen), kann die Ermittlung der Erkennungsleistung sinnvollerweise nur auf Basis der Ergebnisse des 1:2-Vergleichs erfolgen. Bei der Gesichtserkennung gibt es nur ein Merkmal, daher erfolgt die Ermittlung der Erkennungsleistung hierfür auf Basis des 1:1-Vergleichs.

Nach Bereinigung der protokollierten Daten gemäß den Kriterien 1 bis 5 und der Klassifikation gemäß den Kriterien 6 und 7 resultieren die in Tabelle 23 dargestellten Mengen von Verifikationsereignissen (Betätigungen), auf denen sämtliche FRR-Berechnungen beruhen.

System	RefID	Wenignutzer	Sporadische Nutzer	Normalnutzer	Vielnutzer	Alle Nutzer
Cognitec (GE)	1 (ICAO-Bild) <sup>28</sup>	915	3.570	30.605	30.431	65.521
	2 (Template) <sup>29</sup>	947	3.354	7.623	1.645	13.569
BDr/NEC (FE)	3 (ICAO-Bilder)	1.883	7.144	38.700	32.669	80.396
	4 (Template)	1.885	7.167	38.872	32.649	80.573
Dermalog (FE)	5 (ICAO-Bilder)	1.836	8.567	36.534	20.533	67.470
	6 (Template)	1.793	8.555	36.586	20.284	67.218
SD Ind. (IE)	7 (ICAO-Bilder)	1.896	7.715	36.404	20.168	66.034
	8 (Template)	1.747	8.243	38.185	21.109	69.433

**Tabelle 23: Umfang der nutzerklassenspezifischen Datenbasen**

### 7.2.1.2 Verifikationsereignisse zur FAR-Ermittlung

Zur Ermittlung von statistisch aussagekräftigen Falschakzeptanzraten ist eine große Anzahl von Verifikationen unberechtigter Personen erforderlich. Dies erfolgte in einem Offline-Verfahren nach Abschluss des Feldtests.

Von jeder Person wurde je Systemgruppe ein repräsentatives Live-Bild des primären Merkmals aus dem Feldtest oder dem TeachIn ausgewählt, das für die betreffende Person zu einer positiven Verifikation geführt hat. Dieses wurde in einem Batchlauf vom jeweiligen System gegen jeweils das primäre Merkmal beider Referenzbasen (ICAO-Bild und Template aus Live-Enrolment) aller Testteilnehmer verglichen. Diese Vorgehensweise entspricht einer technischen Evaluierung gemäß [ISO\_BPTR]. Der Umfang der resultierenden Datenbasis ist in Tabelle 24 dargestellt.

<sup>28</sup> Nur Testzyklus 2 (siehe Abschnitt 6.2.2.1)

<sup>29</sup> Nur Testzyklus 1 (siehe Abschnitt 6.2.2.1)

System	RefID	Anzahl Input Bilder aus Feldtest/TeachIn	Anzahl Referenzen für Vergleich	Anzahl Ereignisse für FAR-Berechnung
Cognitec (GE)	1 (ICAO-Bild)	1.957	2.070	4.049.040
	2 (Template)	1.957	2.017	3.945.359
BDr/NEC (FE)	3 (ICAO-Bilder)	1.953	2.047	3.995.856
	4 (Template)	1.953	2.051	4.003.660
Dermalog (FE)	5 (ICAO-Bilder)	1.785	2.047	3.652.126
	6 (Template)	1.785	2.032	3.625.363
SD Ind. (IE)	7 (ICAO-Bilder)	2.017	1.927	3.884.850
	8 (Template)	2.017	2.071	4.175.191

**Tabelle 24: Datenbasis für FAR-Ermittlung**

Diese Ergebnisse zeigen, dass die für die FAR-Ermittlung herangezogene Datenbasis hinreichend groß ist, um signifikant eine sehr kleine Schwankungstoleranz der FAR bei dem jeweiligen Schwellwert zu gewährleisten.

### Ursachen für Falschakzeptanzen

Es erfolgte eine Analyse der Live-Bilder und Referenzen für die Verifikationen, die zu hohen Matchscores geführt haben. Ziel dieser Untersuchung war es, Ursachen zu ermitteln, die zu Falschakzeptanzen führen.

- **Cognitec**

Beim Cognitec-Algorithmus sind Falschakzeptanzen mit hohen Übereinstimmungswerten nahezu ausschließlich auf Personen zurückzuführen, die Brillen mit einem markanten dunklen Rand tragen. Bei der Referenzbasis ICAO-Bild sind unter den einhundert Verifikationen Unberechtigter mit den höchsten Matchscores sowohl bei den Referenzbildern als auch bei den Live-Bildern ausschließlich Personen mit einem entsprechendem Brillentyp (unter den 660 höchsten Matchscores bei Verifikationen Unberechtigter sind sowohl bei den Referenz- als auch bei den Live-Bildern jeweils etwa 580 Brillenträger).<sup>30</sup> Auch bei Verwendung des System-Templates als Referenzbasis ist dieses Problem in etwa gleichem Maße vorhanden (97 der 100 Verifikationen Unberechtigter mit den höchsten Matchscores betreffen Personen mit diesem Brillentyp).

- **Bundesdruckerei/NEC**

Verifikationen Unberechtigter, die zu hohen Matchscores führen, sind zum Teil auf schlechte Merkmalsaufnahmen zurückzuführen (wenig Kontrast, kleiner Abdruck, Zentrum des Fingerabdrucks nicht im Bild oder nur am Rand etc.). Diese Beobachtung lässt sich jedoch nicht verallgemeinern, da in etwa gleichem Maße auch mit qualitativ guten Aufnahmen ähnliche Matchscores erreicht werden.

Für dieses System ist anzumerken, dass alle Matchscores, die bei den Verifikationen Unberechtigter erzielt wurden, unterhalb des im Feldtest konfigurierten Schwellwertes lagen.

<sup>30</sup> Auch bei einer von Cognitec bereitgestellten Weiterentwicklung des Algorithmus (Stand September 2004) ist das Problem im gleichen Ausmaß vorhanden.

- **Dermalog**

Bei Dermalog lassen sich anhand der Bilder keine möglichen Ursachen für Falschakzeptanzen ermitteln.

- **SD Industries**

Bei SD Industries lassen sich anhand der Bilder keine möglichen Ursachen für Falschakzeptanzen ermitteln.

## 7.2.2 Failure-To-Enrol Rate (FTE)

Vor dem Hintergrund des Zielszenarios ist es von besonderer Bedeutung, dass alle Personen eingelernt werden können. Bei biometrischen Systemen ist dies nicht immer gegeben. Der Anteil der Population, für das ein System nicht in der Lage ist, den Enrolment-Prozess abzuschließen, wird in der FTE (Failure-To-Enrol Rate) ausgedrückt. Dieser Fehler beinhaltet folgende Ursachen:

- Die Person ist nicht in der Lage, das erforderliche biometrische Merkmal zu präsentieren.
- Das System kann keine Merkmalsaufnahme mit ausreichender Qualität beim Enrolment generieren.
- Die Testverifikation beim Enrolment schlägt fehl (Testperson wird nicht erkannt).

An diesen Ursachen ist erkennbar, dass die Enrolment-Policy wesentlichen Einfluss auf die FTE hat. Wesentliche Parameter hierzu sind die Anzahl der Enrolment-Versuche, die Festlegung zu Testverifikationen sowie Qualitätsschwellwerte für die Merkmalsaufnahme. Die entsprechenden Festlegungen können den Abschnitten 6.1.6 und 6.2.1.8 entnommen werden.

In BioP II wurden zur Ermittlung der FTE für das Enrolment an den biometrischen Systemen (Generierung herstellerepezifischer Templates) sowohl die entsprechenden Enrolment-Protokollierungen der Systeme als auch Checklisten des Enrolment-Personals berücksichtigt.

Die Ermittlung der FTE für die bereitgestellten ICAO-Bilder erfolgte auf Basis der bei einer Verifikation des betreffenden Teilnehmers generierten Protokolleinträge. Hintergrund ist, dass für die ICAO-Bilder bei jeder Verifikation zunächst für die innerhalb der Referenzdatenbank vorliegende Bilddatei ein Enrolment durchgeführt wurde.

Die Ergebnisse sind in Tabelle 25 zusammengefasst. Dabei ist zu beachten, dass das Enrolment für eine Person bereits als gescheitert gilt, wenn eins der geforderten biometrischen Merkmale für ein System nicht enrolt werden kann (zum Beispiel nicht die linke Iris oder der linke Zeige- oder Mittelfinger).

System	RefID	FTE [%]	Teilnehmer	Enrolment gescheitert
Cognitec (GE)	1 (ICAO-Bild)	0	1627	0
	2 (Template)	0	2077	0
BDr/NEC (FE)	3 (ICAO-Bild)	0,25	1620	4
	4 (Template)	0,99	2018	20
Dermalog (FE)	5 (ICAO-Bild)	0	1625	0
	6 (Template)	0,40	2019	8
SDI (IE)	7 (ICAO-Bild)	0	1540	0
	8 (Template)	0,89	2017	18

**Tabelle 25: FTE<sup>31</sup>**

### 7.2.2.1 FTE-Klassifikation

Auf Basis der Enrolment-Protokolle, Erfahrungen des Enrolment-Personals sowie während des Enrolments aufgezeichneter Protokolldaten und Bildaufnahmen wurden wesentliche Ursachen für gescheiterte Enrolments klassifiziert. Diese sind nachfolgend aufgelistet.

- **Gesicht:** Es gab keine gescheiterten Enrolments.
- **Finger:** Bei der Aufnahme von Fingerabdrücken wurden folgende Ursachen festgestellt, die ein erfolgreiches Enrolment verhinderten:
  - Feuchte Finger: Diese Ursache korreliert zum Teil mit einer möglichen Aufregung beim Enrolment-Vorgang. Es wurde festgestellt, dass für Personen mit auffallend feuchten Fingern beim Enrolment während des Feldtests kontrastreichere (und damit weniger feuchte Finger) Bilder aufgezeichnet wurden.
  - Trockene Finger
  - Gestörte bzw. schwierige Struktur der Finger
  - Fehler durch Enrolment-Personal

Für einzelne Personen war die Ursache nicht ermittelbar.

- **Iris:** Bei der Iris-Aufnahme wurden folgende Ursachen festgestellt, die ein erfolgreiches Enrolment verhinderten:
  - Sehschwäche: Einzelnen Testteilnehmern war es durch eine Sehschwäche nicht möglich, das Merkmal korrekt zu positionieren. Dieser Effekt trat auf, da Brillen während des Enrolments gemäß Herstellerempfehlung abgesetzt werden mussten.
  - Systemdesign: Aufgrund der Einbauhöhe der Kamera des bereitgestellten Systems konnten Personen mit einer Körpergröße unter 1,55 m schwer bis gar nicht enrolt werden.
  - Fehler durch Enrolment-Personal

Für einzelne Personen war die Ursache nicht ermittelbar.

<sup>31</sup> Ursache für die Differenz der Teilnehmerzahlen zwischen ICAO-Bild und Template ist die Tatsache, dass am Feldtest (Grundlage für FTE bei ICAO-Bild) deutlich weniger Personen teilgenommen haben als am Enrolment (Grundlage für FTE bei Template).

### 7.2.3 Erkennungsleistungen

Nachfolgend werden die erzielten Erkennungsleistungen der biometrischen Systeme innerhalb des BioP II Feldtests dargestellt. Dabei wird entsprechend der Motivation aus Abschnitt 7.1.2.2 in Gesamterkennungsleistung, nutzerklassenspezifische Erkennungsleistung sowie Einzelbenutzerstatistik unterschieden.

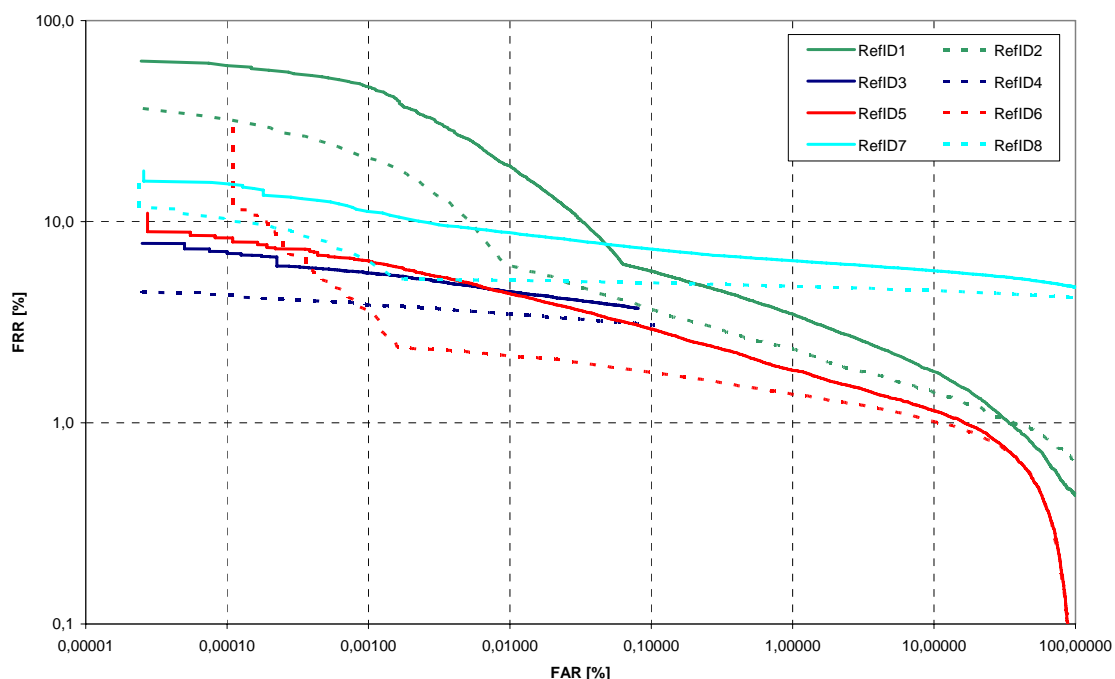
Bei der Einordnung der erzielten Matchscores ist folgendes zu berücksichtigen:

- **Cognitec**
  - Je größer der Matchscore desto besser die Übereinstimmung
  - Wertebereich:  $0 \leq \text{Matchscore} \leq 1$
- **Bundesdruckerei/NEC**
  - Je größer der Matchscore desto besser die Übereinstimmung
  - Wertebereich:  $0 \leq \text{Matchscore} \leq 20.000$
- **Dermalog**
  - Je größer der Matchscore desto besser die Übereinstimmung
  - Wertebereich:  $0 \leq \text{Matchscore} \leq 100$
- **SD Industries**
  - Je kleiner der Matchscore desto besser die Übereinstimmung
  - Wertebereich:  $0 \leq \text{Matchscore} \leq 1$

Bei Bundesdruckerei/NEC kann für verschiedene FAR-Arbeitspunkte keine Toleranzschwelle ermittelt werden. Dies ist darin begründet, dass beim minimal konfigurierbaren Schwellwert die FAR bei 0,081% (RefID3) bzw. 0,11% (RefID4) liegt. Wenn im Folgenden die Erkennungsleistung anhand von festen Arbeitspunkten betrachtet wird, ist dies entsprechend zu berücksichtigen.

#### 7.2.3.1 Gesamterkennungsleistung

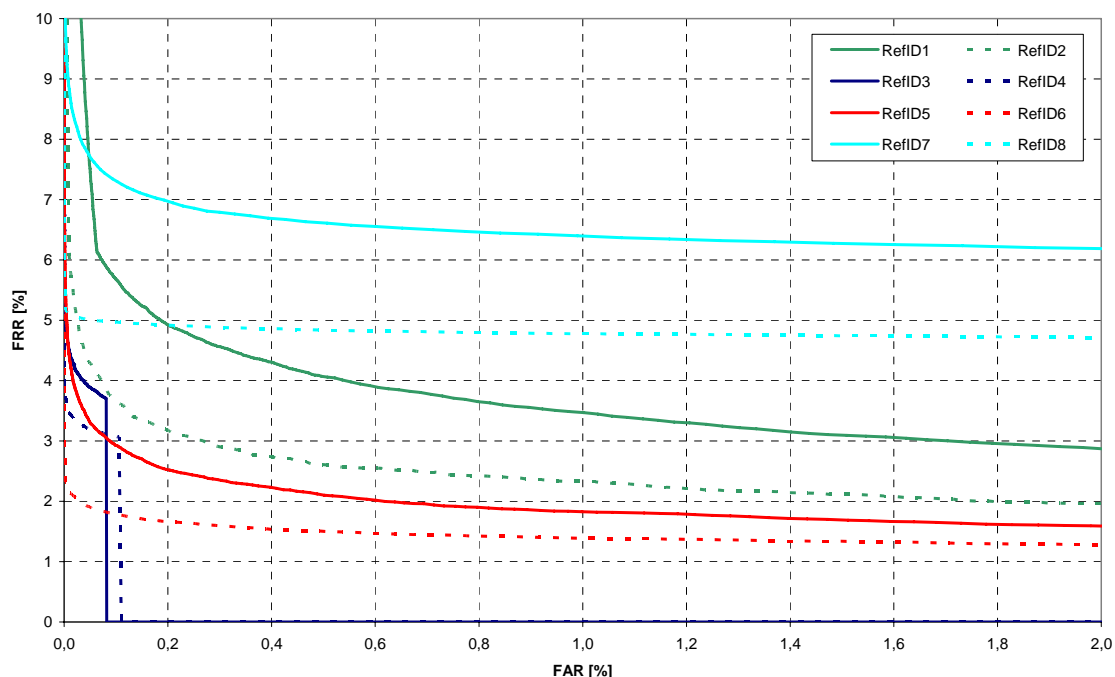
Für die Ermittlung der Gesamterkennungsleistung finden alle regulären Betätigungen Eingang. Dabei ist zu berücksichtigen, dass aufgrund von zum Teil sehr unterschiedlichen Betätigungszahlen der Testteilnehmer einzelne Personen einen stärkeren Einfluss auf die Ergebnisse haben als andere. Um die Ergebnisse der verschiedenen Systeme und Referenzbasen vergleichen zu können, werden diese auf Basis von DET-Kurven dargestellt. Abbildung 57 zeigt die im Feldtest erzielte Performance über den gesamten Arbeitsbereich der Systeme. Um auch die interessanten Bereiche ( $\text{FAR} < 1\%$ ) hinreichend gut darstellen zu können, wurde eine logarithmische Skalierung gewählt.



**Abbildung 57: Vergleich der DET-Kurven (logarithmische Skalierung des Arbeitsbereichs)**

Die „Knicke“ in einzelnen DET-Kurven verdeutlichen den Arbeitspunkt, an dem die betreffende RefID während des Feldtests als Masterreferenz mit entsprechendem Schwellwert konfiguriert war. Dies betrifft RefID1 (Masterreferenz im Testzyklus 2), RefID2 (Masterreferenz im Testzyklus 1), RefID6 und RefID8. Für RefID4, die im System von Bundesdruckerei/NEC als Masterreferenz eingestellt war, zeigt sich nicht das beschriebene Verhalten. Dies liegt daran, dass der im Feldtest konfigurierte Schwellwert zu einer FAR von 0% führt und somit außerhalb des logarithmisch darstellbaren Bereichs liegt.

Da für die Bewertung der biometrischen Systeme Arbeitsbereiche irrelevant sind, bei denen die FAR größer als 2% oder die FRR größer als 10% ist, erfolgt in Abbildung 58 eine Beschränkung auf die entsprechenden Bereiche.



**Abbildung 58: Vergleich der DET-Kurven über den Arbeitsbereich FAR≤2%, FRR≤10%**

Anhand der DET-Kurven wird deutlich, dass es kein einheitliches Ranking der verschiedenen Systeme gibt. Für verschiedene Sicherheitsniveaus (FAR-Arbeitspunkte) ist die Rangfolge der Systeme hinsichtlich der FRR unterschiedlich.

Bei allen Systemen zeigt sich, dass das aus dem Live-Enrolment gewonnene Referenz-Template besser abschneidet als das ICAO-Referenzbild. Hierbei ist jedoch zu beachten, dass (mit Ausnahme von Cognitec) das Template jeweils als Masterreferenz verwendet wurde. Ist das ICAO-Bild als Masterreferenz konfiguriert, ergibt sich für dieses eine Verbesserung der Erkennungsleistung, da das bei der Betätigung aufgenommene Live-Bild hinsichtlich der Masterreferenz optimiert ausgewählt wird.

Weitere Details zur Erkennungsleistung der verschiedenen Systeme und Referenzen werden anhand der Matchscore-Histogramme (Genuine-Impostor-Frequency-Diagramme) sowie der FAR-FRR-Diagramme deutlich. Diese sind in den folgenden Abbildungen dargestellt.

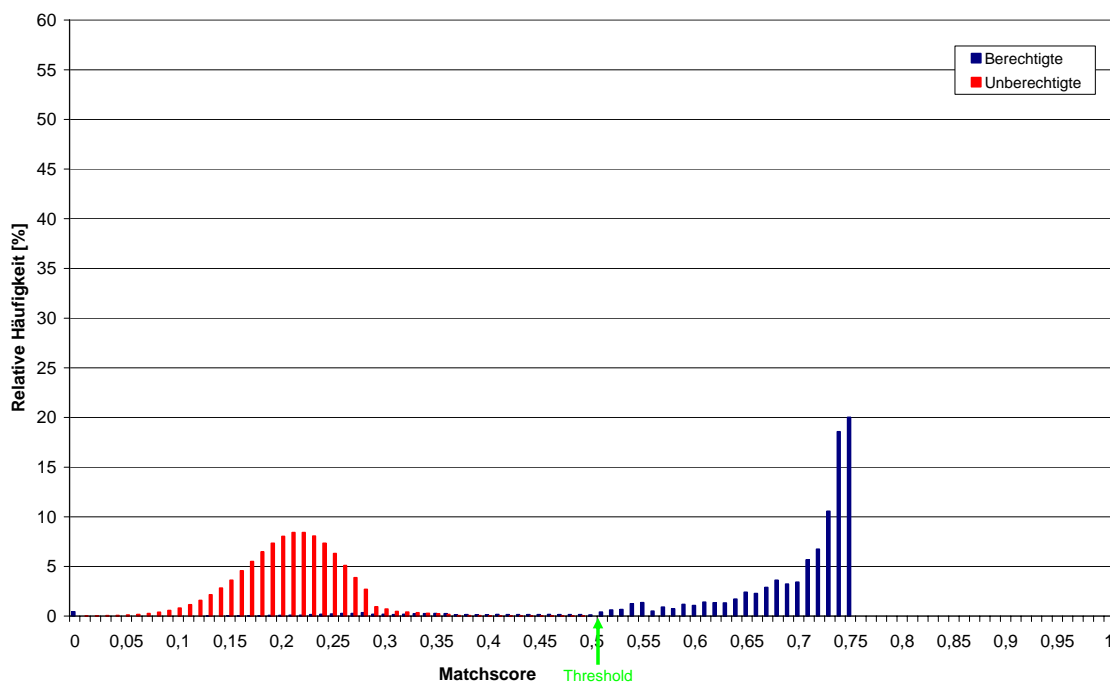


Abbildung 59: Matchscore-Verteilung für RefID1 (Cognitec, ICAO-Bild, Testzyklus 2)

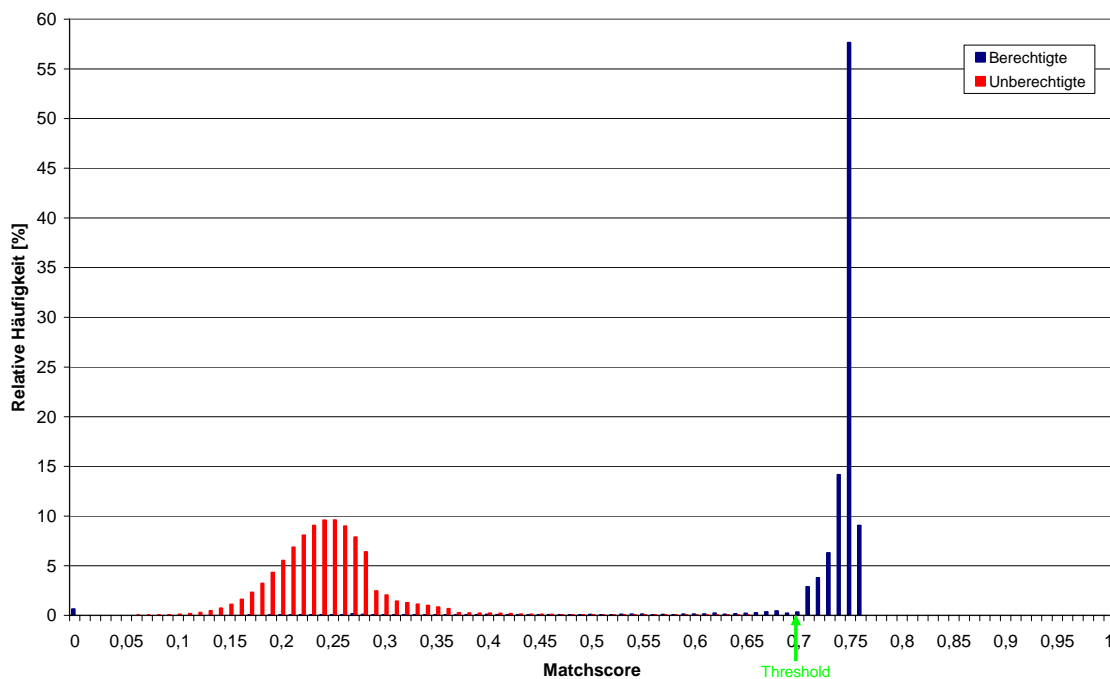


Abbildung 60: Matchscore-Verteilung für RefID2 (Cognitec, Template, Testzyklus 1)

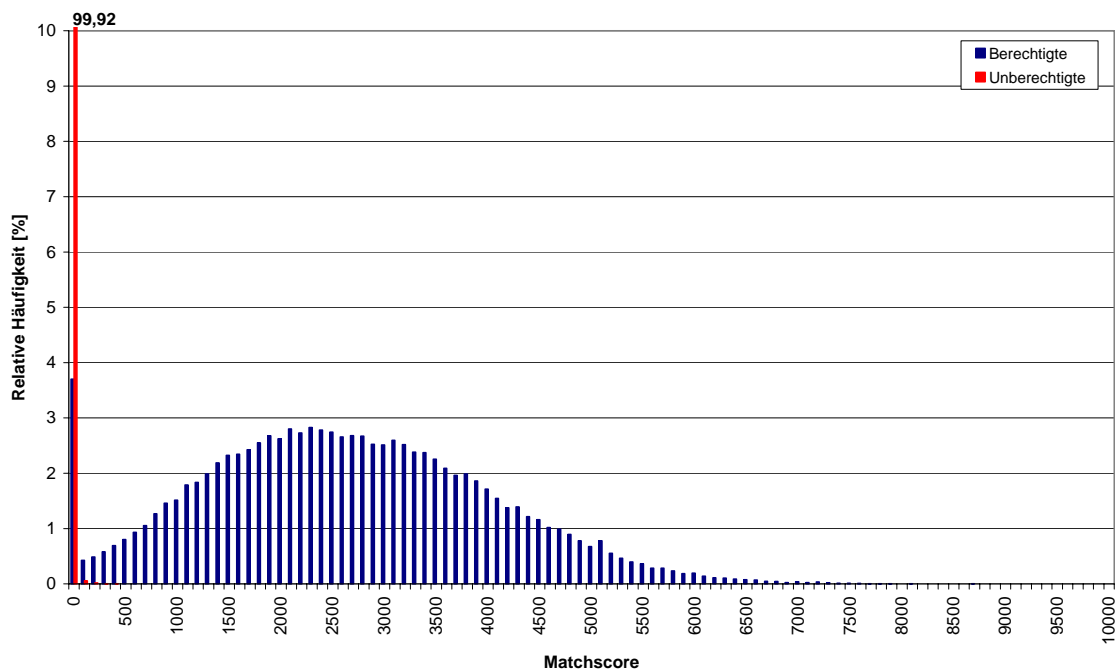


Abbildung 61: Matchscore-Verteilung für RefID3 (BDr/NEC, ICAO-Bilder)

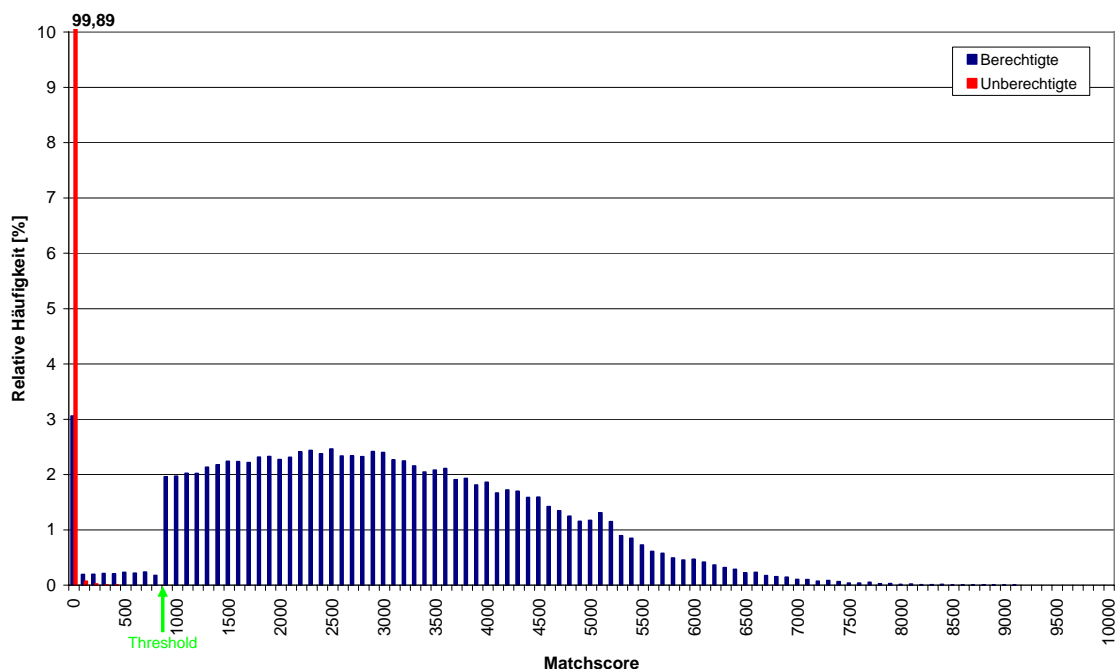


Abbildung 62: Matchscore-Verteilung für RefID4 (BDr/NEC, Template)

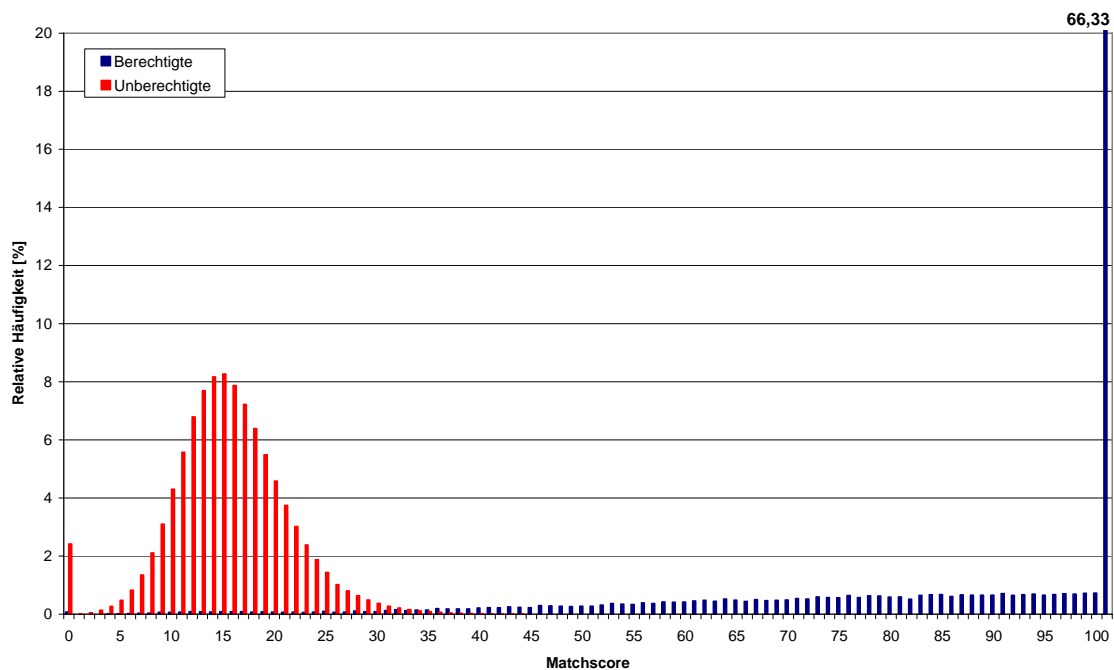


Abbildung 63: Matchscore-Verteilung für RefID5 (Dermalog, ICAO-Bilder)

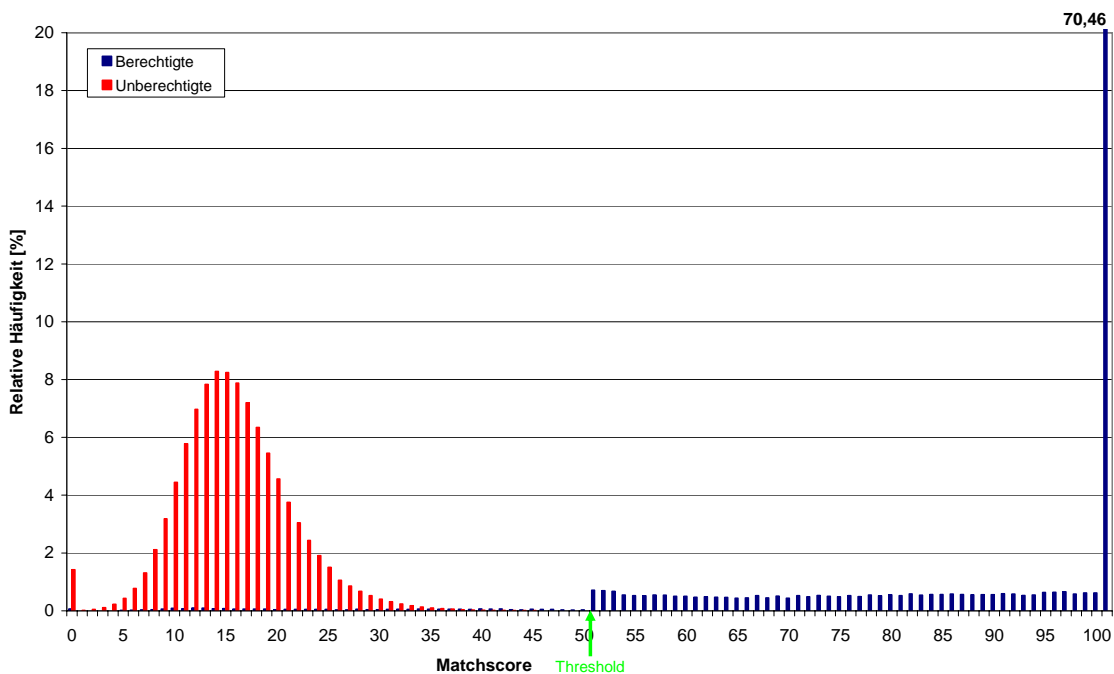


Abbildung 64: Matchscore-Verteilung für RefID6 (Dermalog, Template)

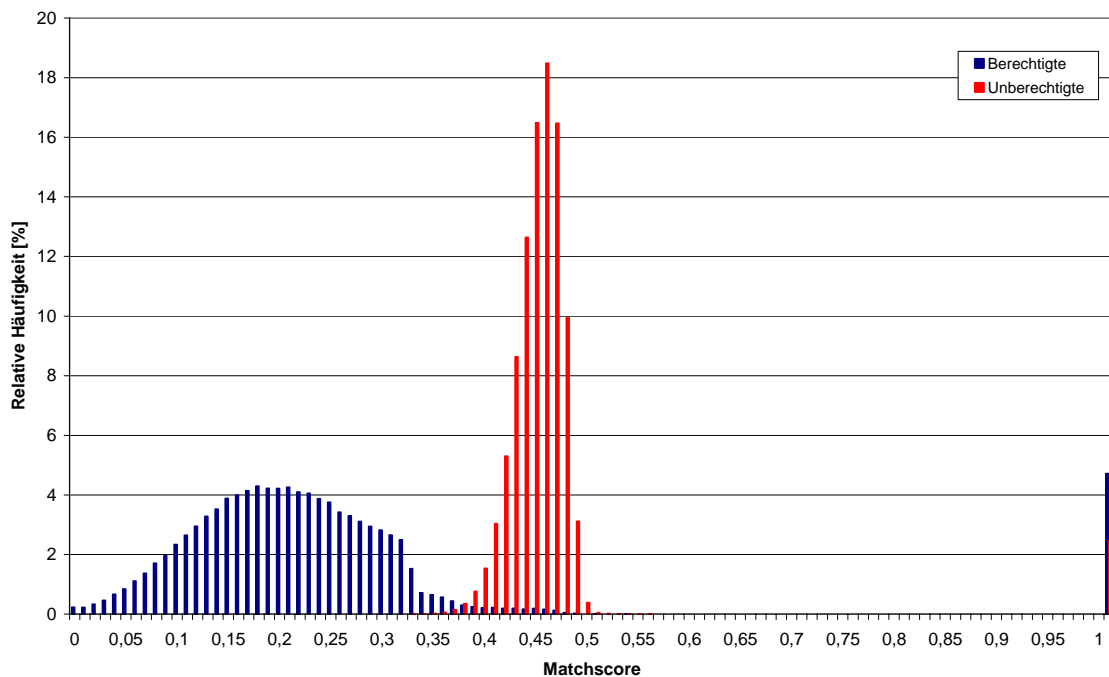


Abbildung 65: Matchscore-Verteilung für RefID7 (SD Industries, ICAO-Bilder)

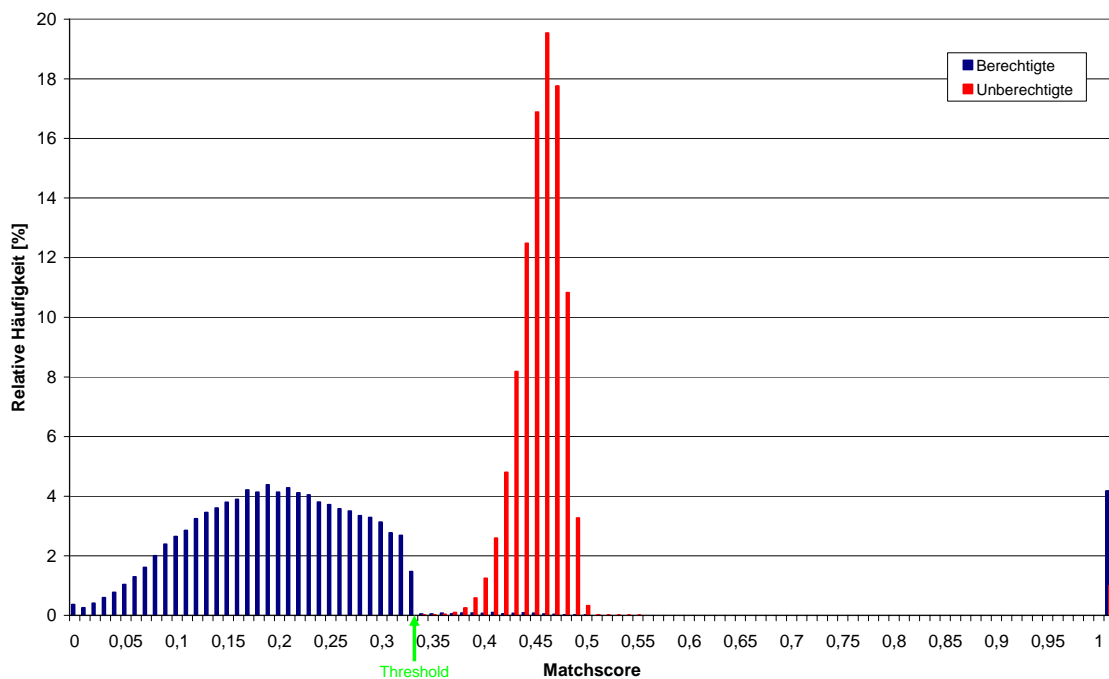


Abbildung 66: Matchscore-Verteilung für RefID8 (SD Industries, Template)

---

Anhand der Genuine-Impostor-Frequency-Diagramme werden einige systemspezifische Details deutlich:

- Bei Bundesdruckerei/NEC (Abbildung 61 und Abbildung 62) wird deutlich, dass der Algorithmus die Verifikationen Unberechtigter nahezu komplett auf den Matchscore 0 abbildet.
- Bei Dermalog (Abbildung 63 und Abbildung 64) wird deutlich, dass der Algorithmus die Verifikationen Berechtigter zu mehr als zwei Drittel auf den Matchscore 100 abbildet.
- Bei SD Industries (Abbildung 65 und Abbildung 66) fällt auf, dass der Algorithmus sowohl bei den Verifikationen Berechtigter als auch bei den Verifikationen Unberechtigter etwa 1% bis 5% auf den Matchscore 1 abgebildet werden. Für diese Fälle gibt es zwei Ursachen:
  - Es konnte kein hinreichend gutes Irisbild aufgenommen werden.
  - Das aus dem Live-Bild gewonnene Template sowie das Referenz-Template (bzw. das aus dem Referenzbild generierte Template) verfügen nicht über genügend auswertbare Bits.

Bei allen Systemen ist für die Masterreferenz in der Verteilung der Matchscores Berechtigter ein „Sprung“ beim jeweils konfigurierten Schwellwert zu erkennen. Ursache hierfür ist, dass bis zum Erreichen eines Timeout iterativ Merkmalsaufnahmen erfolgen, bis für die Masterreferenz die Toleranzschwelle überschritten wird.

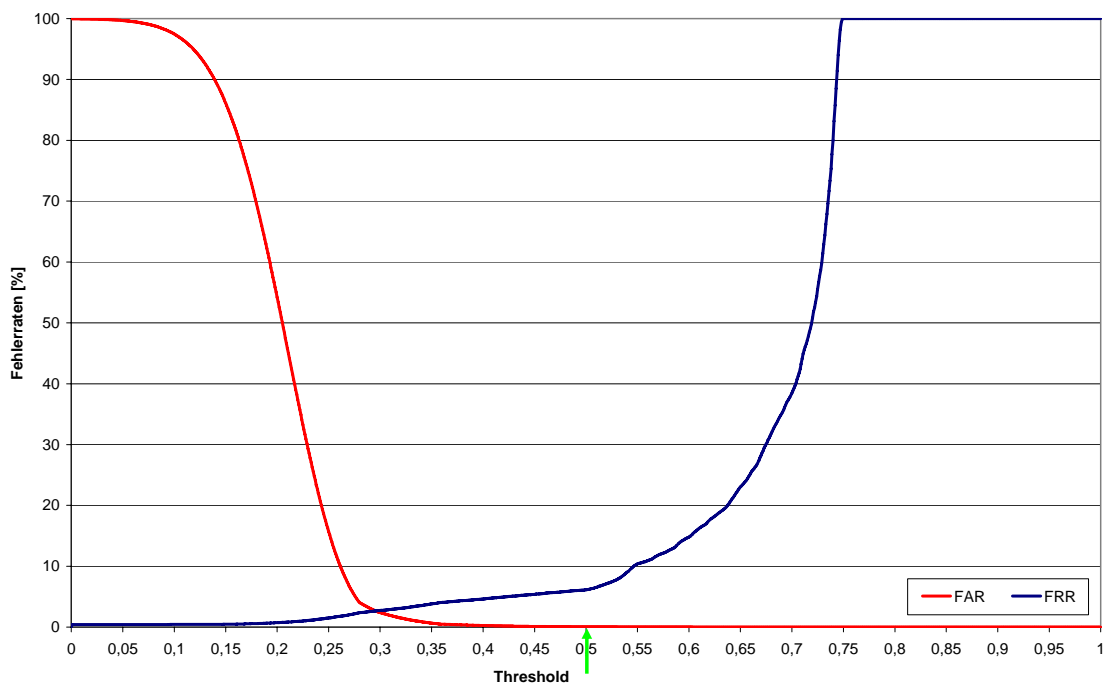


Abbildung 67: FAR-FRR-Kurve für RefID1 (Cognitec, ICAO-Bild, Testzyklus 2)

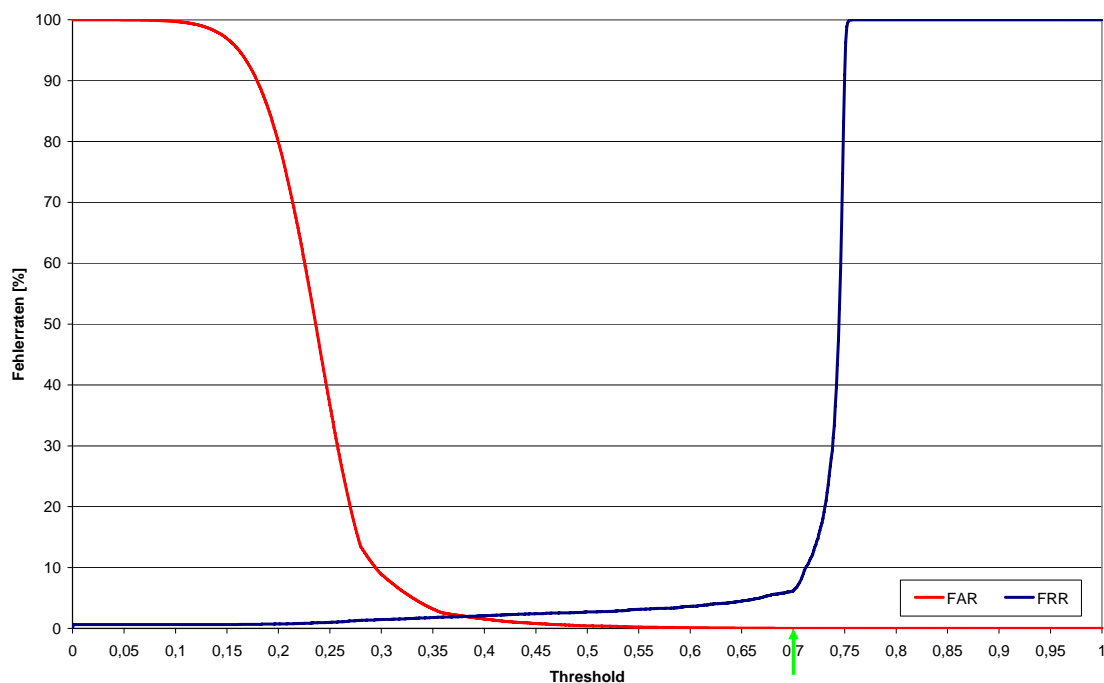


Abbildung 68: FAR-FRR-Kurve für RefID2 (Cognitec, Template, Testzyklus 1)

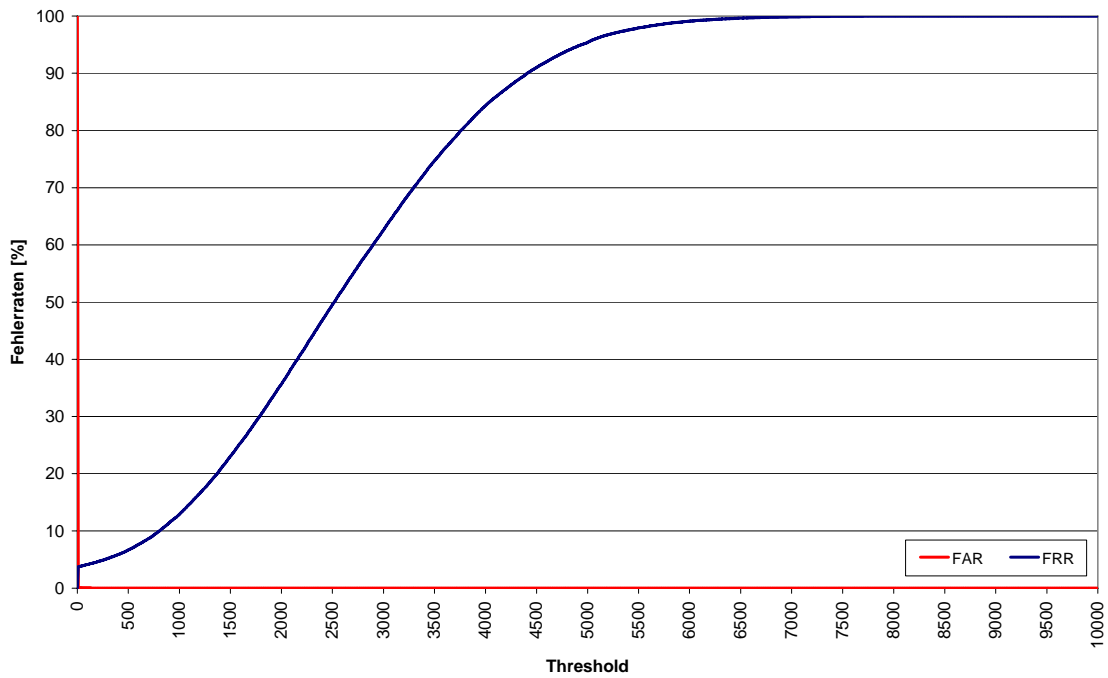


Abbildung 69: FAR-FRR-Kurve für RefID3 (BDr/NEC, ICAO-Bilder)

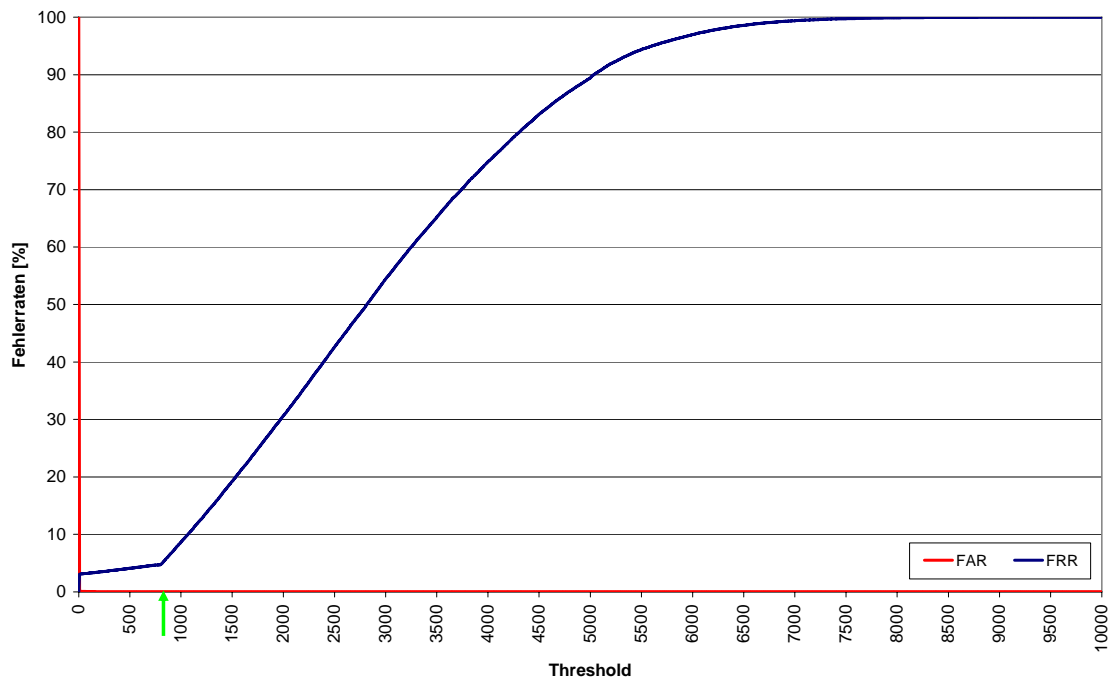


Abbildung 70: FAR-FRR-Kurve für RefID4 (BDr/NEC, Template)

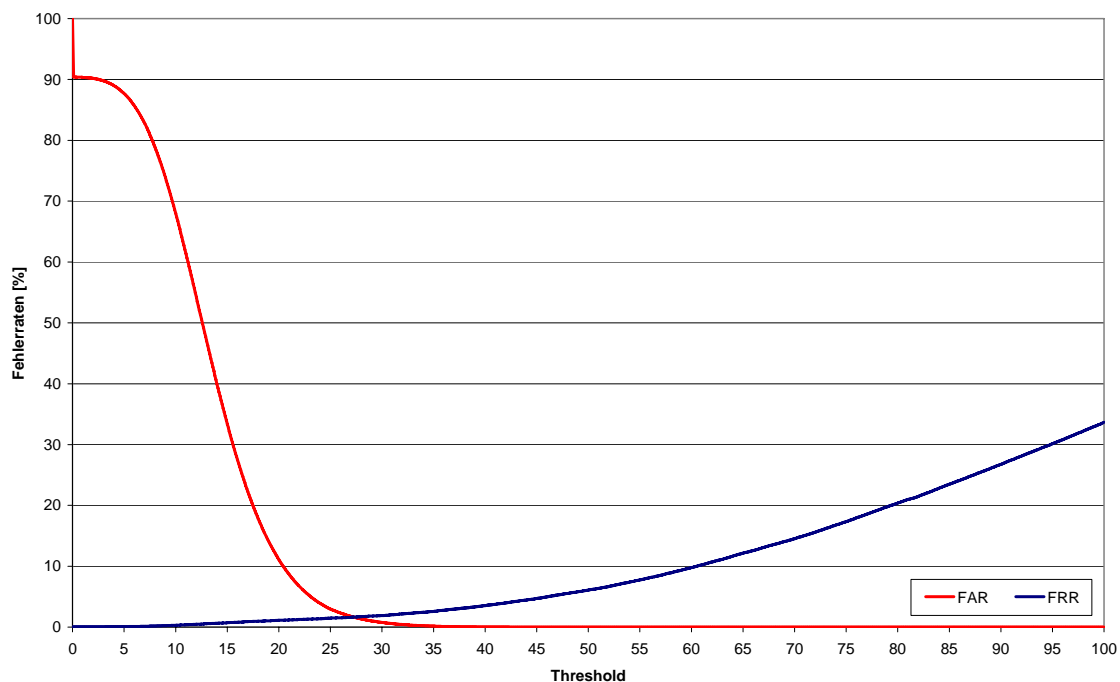


Abbildung 71: FAR-FRR-Kurve für RefID5 (Dermalog, ICAO-Bilder)

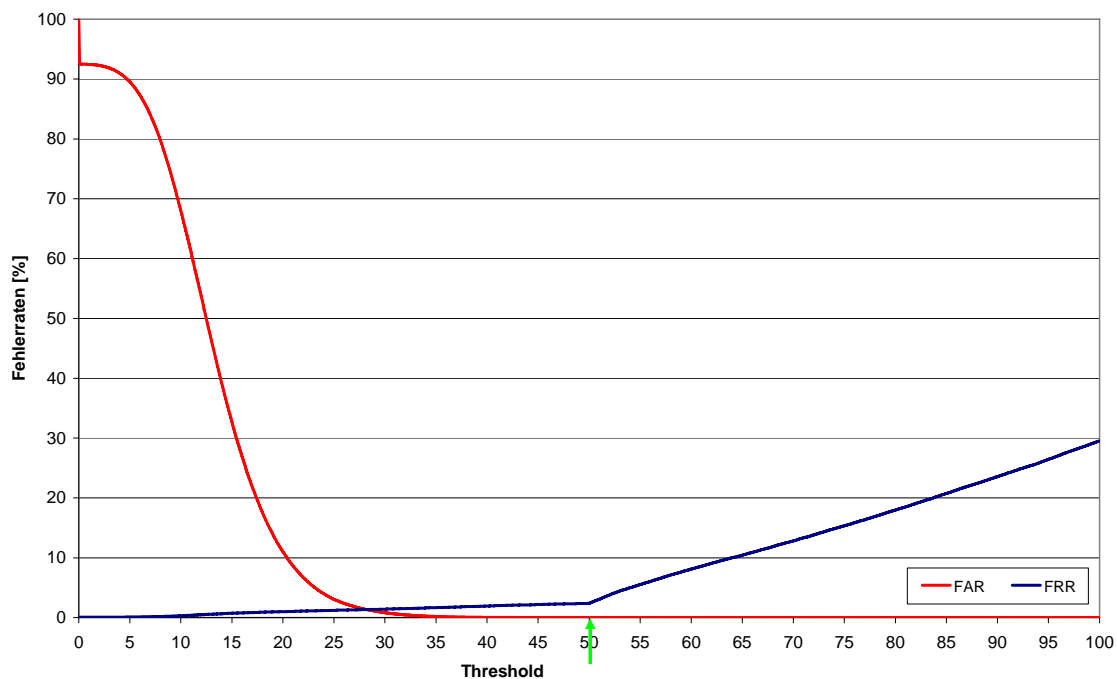


Abbildung 72: FAR-FRR-Kurve für RefID6 (Dermalog, Template)

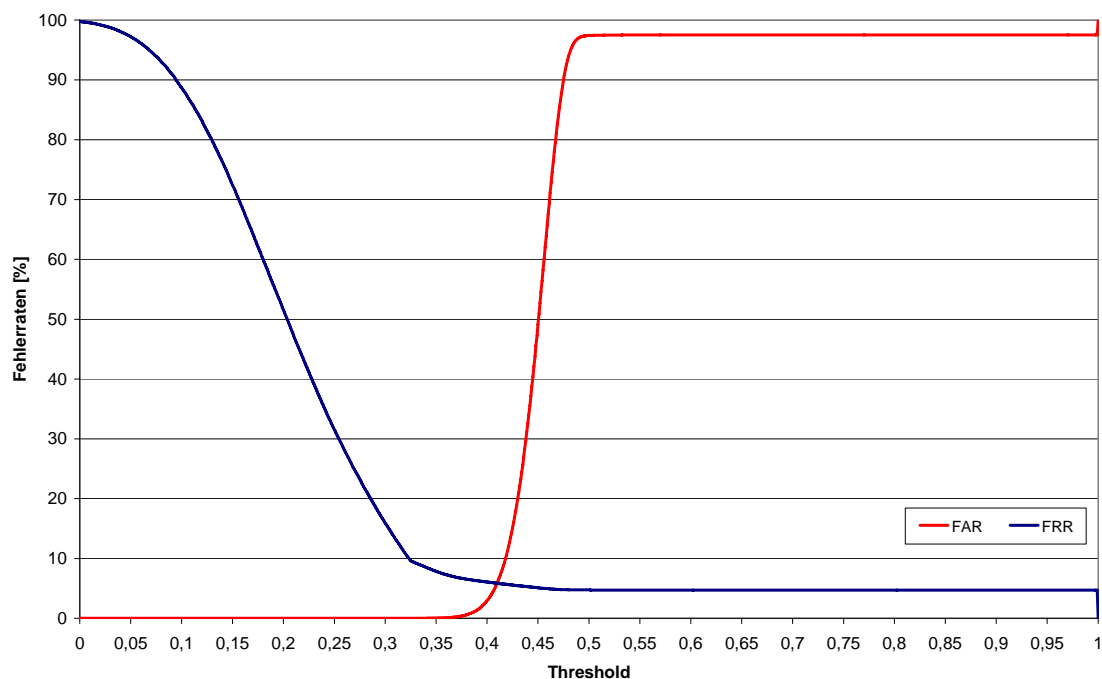


Abbildung 73: FAR-FRR-Kurve für RefID7 (SD Industries, ICAO-Bilder)

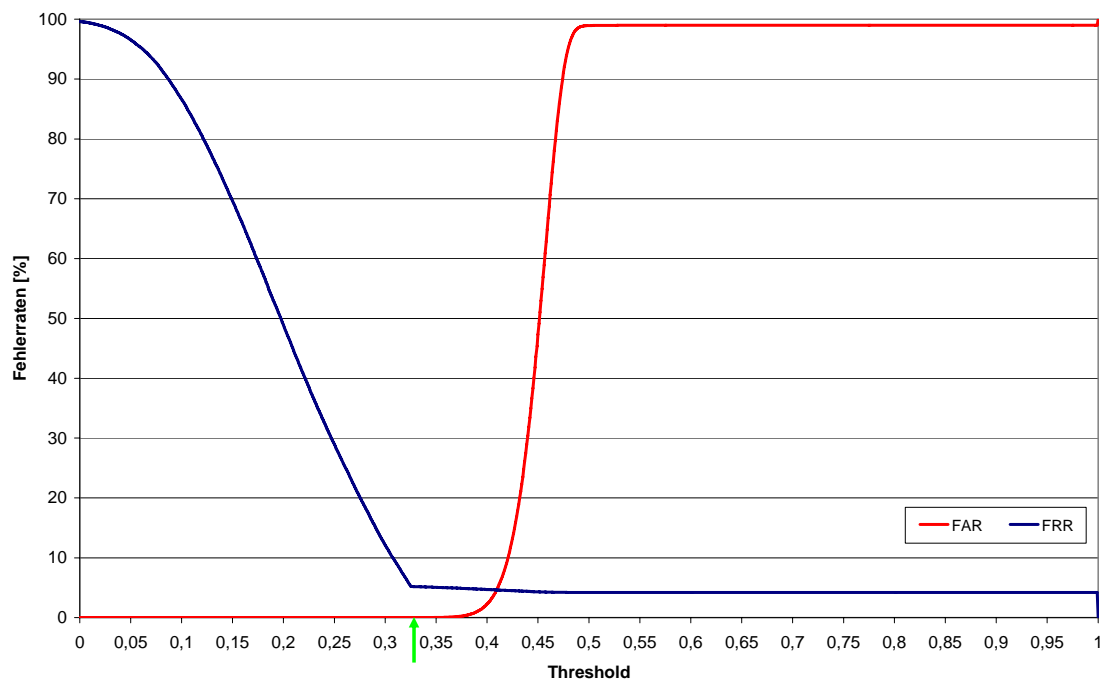


Abbildung 74: FAR-FRR-Kurve für RefID8 (SD Industries, Template)

Anhand der FAR-FRR-Diagramme werden einige systemspezifische Details deutlich:

- Der Algorithmus von Bundesdruckerei/NEC (Abbildung 69 und Abbildung 70) zeigt eine nahezu ideale FAR. Da die Matchscores bei Verifikationen Unberechtigter fast immer 0 sind, fällt die FAR schon beim minimal konfigurierbaren Schwellwert auf 0,081% (RefID3) bzw. 0,11% (RefID4).
- Bei Dermalog (Abbildung 71 und Abbildung 72) nimmt die FRR nicht den typischen Maximalwert von 100% an. Da etwa bei zwei Drittel der Verifikationen Berechtigter der maximal mögliche Matchscore 100 erreicht wird, kann selbst beim Schwellwert 100 die FRR nicht über 35% steigen.

Bei allen Systemen ist für die Masterreferenz in der FRR-Kurve ein „Knick“ beim jeweils konfigurierten Schwellwert zu erkennen.<sup>32</sup>

Aus den Verifikationen Unberechtigter resultierten für die zu betrachtenden Arbeitspunkte die in Tabelle 26 angegebenen Schwellwerte.

RefID	FAR=0,001%	FAR=0,01%	FAR=0,1%	FAR=1%
1	0,715	0,630	0,467	0,336
2	0,732	0,697	0,604	0,433
3	357	176	-	-
4	397	214	6	-
5	51,07	43,61	36,76	29,10
6	51,76	44,40	37,12	29,29
7	0,318	0,336	0,360	0,386
8	0,321	0,343	0,365	0,390

**Tabelle 26: Schwellwerte für zu betrachtende Sicherheitsniveaus**

Da bei Bundesdruckerei/NEC nahezu alle unberechtigten Versuche auf den Matchscore 0 abgebildet werden, ist eine Toleranzschwelle für FAR=0,1% (bei RefID3) und FAR=1% (bei RefID3 und RefID4) nicht ermittelbar. Im Folgenden wird daher für Betrachtungen an diesen Arbeitspunkten der Schwellwert 1 zugrunde gelegt. Dieser führt bei RefID3 zu FAR=0,081% und bei RefID4 zu FAR=0,106%.

Um zu zeigen, wie sich die Erkennungsleistung der Systeme über den Verlauf des Feldtests entwickelt hat, wird für jeden Tag die bis dahin erreichte FRR berechnet. Basis ist der Arbeitspunkt FAR=0,1%. Das Ergebnis ist in Abbildung 75 dargestellt. Typischerweise ist bei biometrischen Systemen zunächst eine Gewöhnungsphase zu beobachten, in der die Falschrückweisungsraten deutlich zurückgehen, und anschließend ein relativ stabiler Verlauf.

<sup>32</sup> Ursache hierfür ist, dass bis zum Erreichen eines Timeout iterativ Merkmalsaufnahmen erfolgen, bis für die Masterreferenz die Toleranzschwelle überschritten wird.

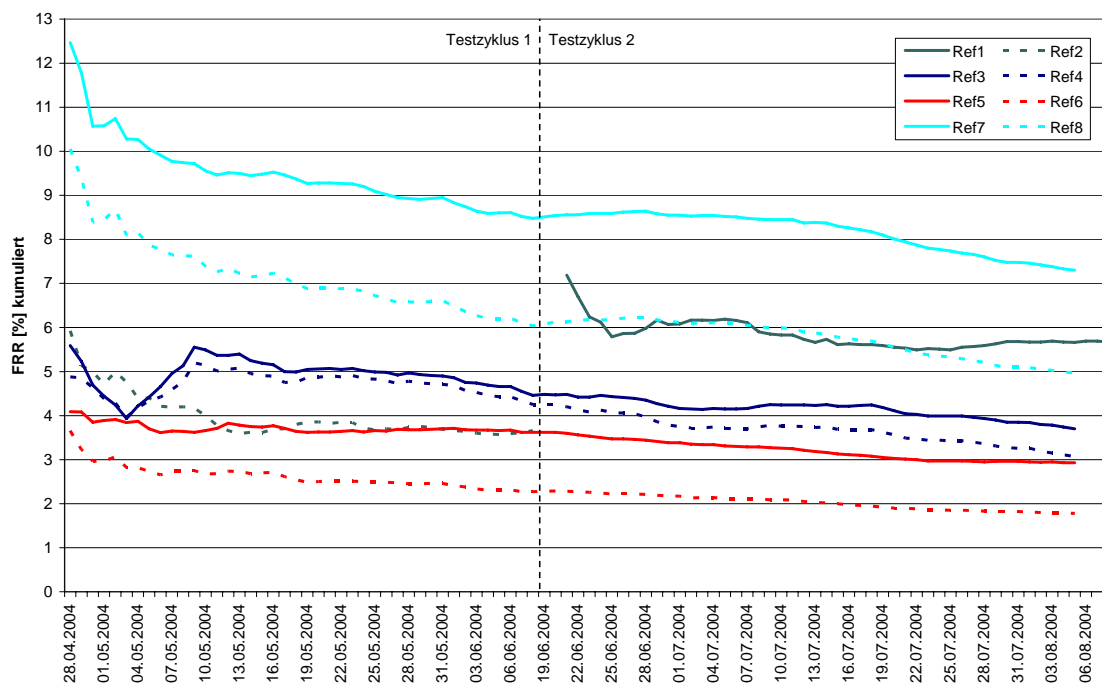


Abbildung 75: FRR im zeitlichen Verlauf (FAR=0,1%)<sup>33</sup>

Bei Cognitec ist für beide Zyklen die erwartete Charakteristik sichtbar. Nach einem Abfallen der FRRs in den ersten zwei Wochen verlaufen die Kurven danach relativ stabil.

Für die anderen drei Systeme ist jedoch selbst zum Ende des Feldtests immer noch ein Sinken der FRR zu beobachten. Dies legt die Vermutung nahe, dass für Fingerabdruck- und Iriserkennung das Training für eine gute Präsentation des Merkmals schwieriger und langwieriger ist als für die Gesichtserkennung. Dennoch erreichen die FE-Systeme bereits von Beginn an eine vergleichsweise niedrige Rückweisungsrate.

Während des Feldtests wurden die Systeme bei festen Schwellwerten für die jeweils konfigurierte Masterreferenz betrieben. Dies führte zu den in Tabelle 27 angegebenen Fehlerraten, wie sie sich für die Testteilnehmer darstellten.

System	RefID	Schwellwert	FRR [%]	FAR [%]
Cognitec (Zyklus 1)	2	0,7	6,13	0,0092
Cognitec (Zyklus 2)	1	0,5	6,13	0,0637
BDr/NEC	4	800	4,73	0
Dermalog	6	50	2,37	0,0016
SD Industries	8	0,33	5,17	0,0023

Tabelle 27: Fehlerraten für Feldtest-Systemkonfiguration

<sup>33</sup> Für RefID3 ist der Arbeitspunkt FAR=0,081%.

### 7.2.3.2 Nutzerklassenspezifische Erkennungsleistung

Die Anzahl der Betätigungen an den biometrischen Systemen unterscheidet sich sehr stark unter den Testteilnehmern. Während einige Benutzer nur äußerst selten die Systeme bedienten, wurden diese von anderen mehrmals pro Tag bedient. Zur Ermittlung entsprechender Einflüsse auf die Erkennungsleistung wurden in Abschnitt 6.1.1.2 Nutzerklassen definiert, für die jeweils in Abbildung 76 die zugehörigen Rückweisungsdaten beim Arbeitspunkt FAR=0,1% angegeben sind.

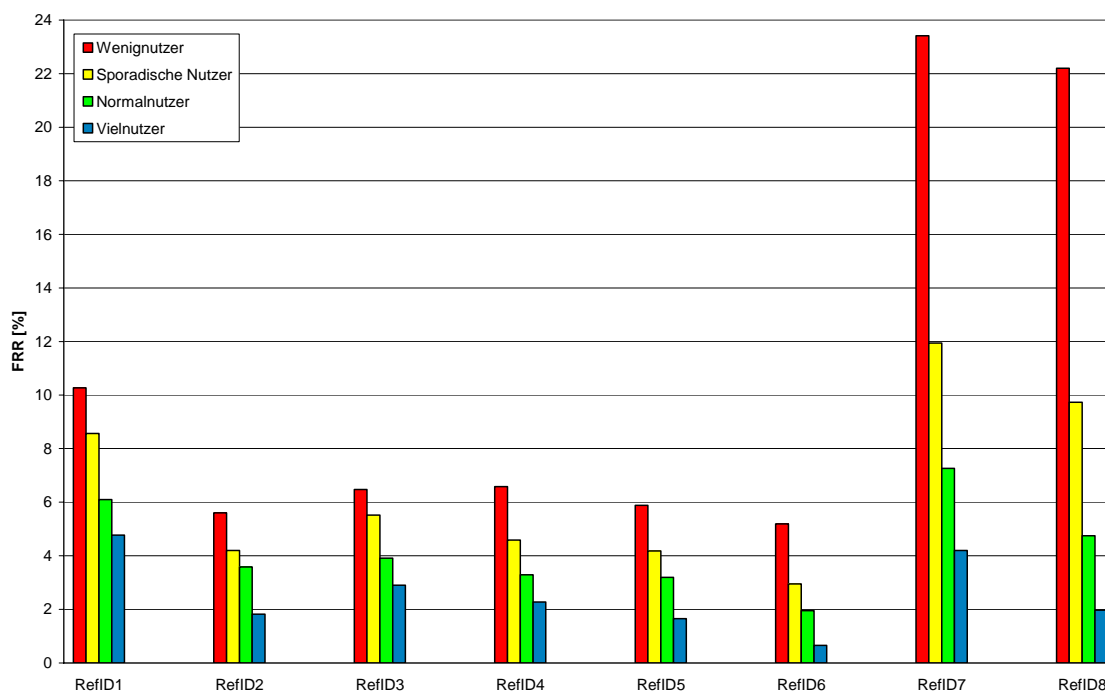


Abbildung 76: FRR nutzerklassenspezifisch (FAR=0,1%)<sup>34</sup>

Bei allen Systemen wird deutlich, dass mit zunehmender Geübtheit der Teilnehmer die FRR deutlich sinkt. Insbesondere bei SD Industries verursachen Wenignutzer und sporadische Nutzer sehr hohe Rückweisungsdaten. Dies legt den Schluss nahe, dass für eine geeignete Merkmalspräsentation ein hohes Maß an Übung erforderlich ist.

### 7.2.3.3 Einzelbenutzerstatistik

Die im Feldtest ermittelten Verifikationsergebnisse unterscheiden sich zum Teil sehr stark zwischen den Benutzern. Um dies näher untersuchen zu können, wird für jeden einzelnen Teilnehmer die FRR ermittelt. Die Ergebnisse sind in den folgenden Abbildungen dargestellt. Dort werden für jede RefID die erreichten Einzelbenutzer-FRRs absteigend angeordnet. Die Anzahl der Teilnehmer, für die Werte in das Diagramm eingegangen sind, entspricht dem maximalen Skalenwert der Abszisse. Der rechte Bereich des Diagramms beinhaltet somit die Teilnehmer, für die FRR=0% gilt.

Neben der Gesamtdarstellung der Einzelbenutzer-FRRs ist im Folgenden auch jeweils eine nutzerklassenspezifische Darstellung der Einzelbenutzer-FRRs angegeben. Auch hier repräsentieren die Bereiche, in denen keine Werte sichtbar sind, Teilnehmer mit FRR=0% in der jeweiligen Nutzerklasse.

<sup>34</sup> Für RefID3 ist der Arbeitspunkt FAR=0,081%.

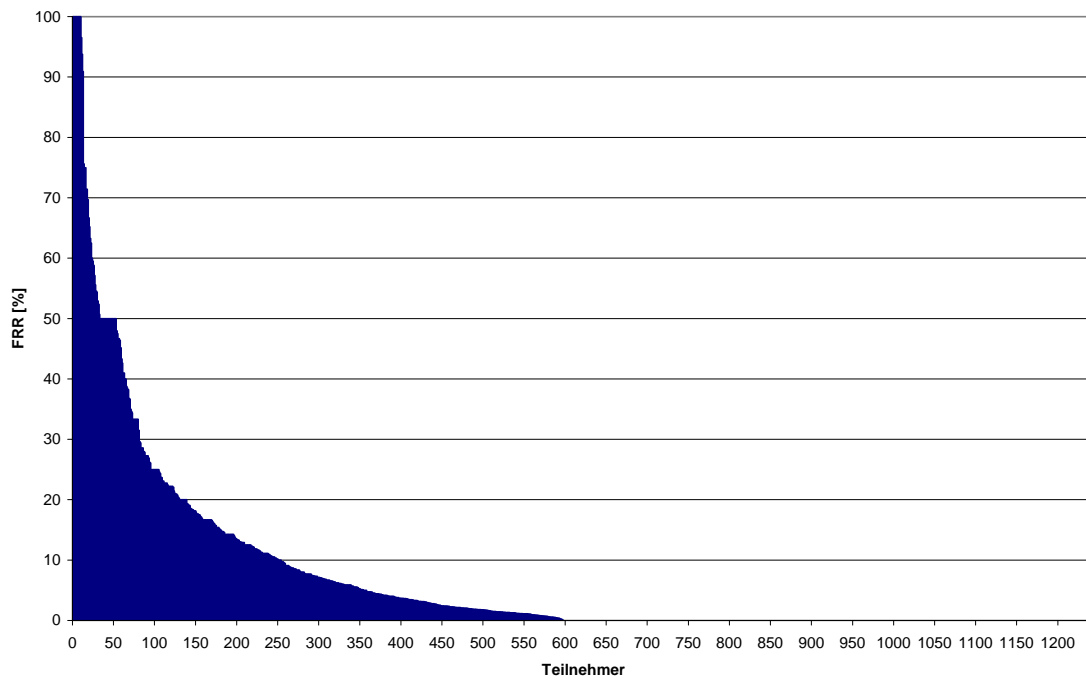


Abbildung 77: Einzelbenutzerstatistik RefID1 (Cognitec, ICAO-Bild, Zyklus 2), FAR=0,1%

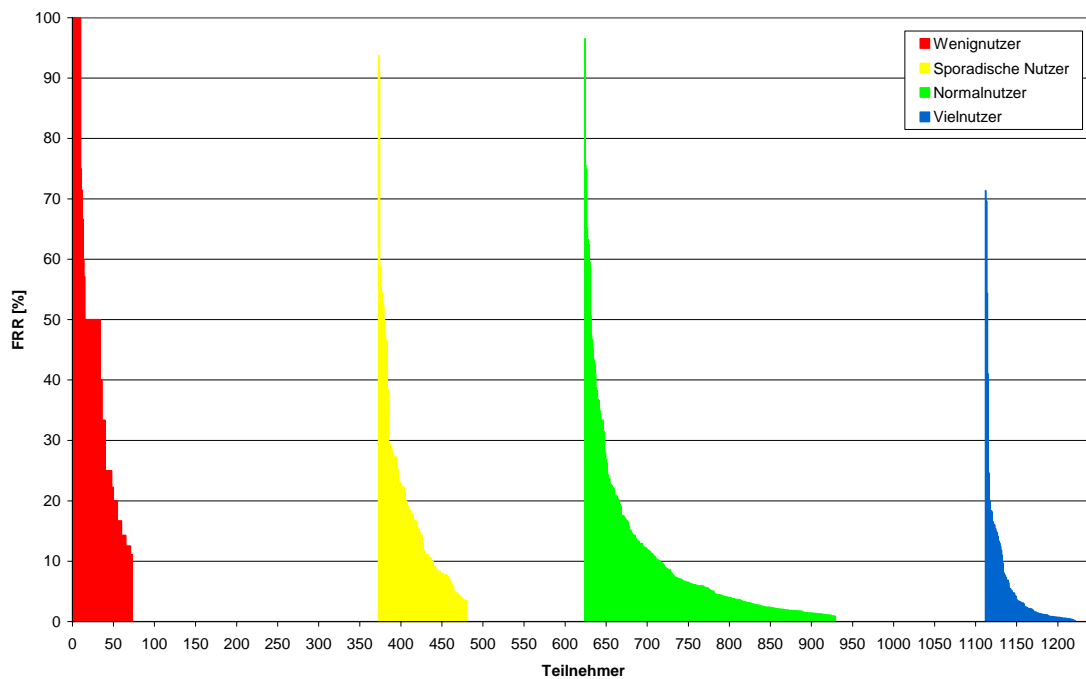


Abbildung 78: Nutzerklassenspezifische Einzelbenutzerstatistik RefID1 (Cognitec, ICAO-Bild, Zyklus 2), FAR=0,1%

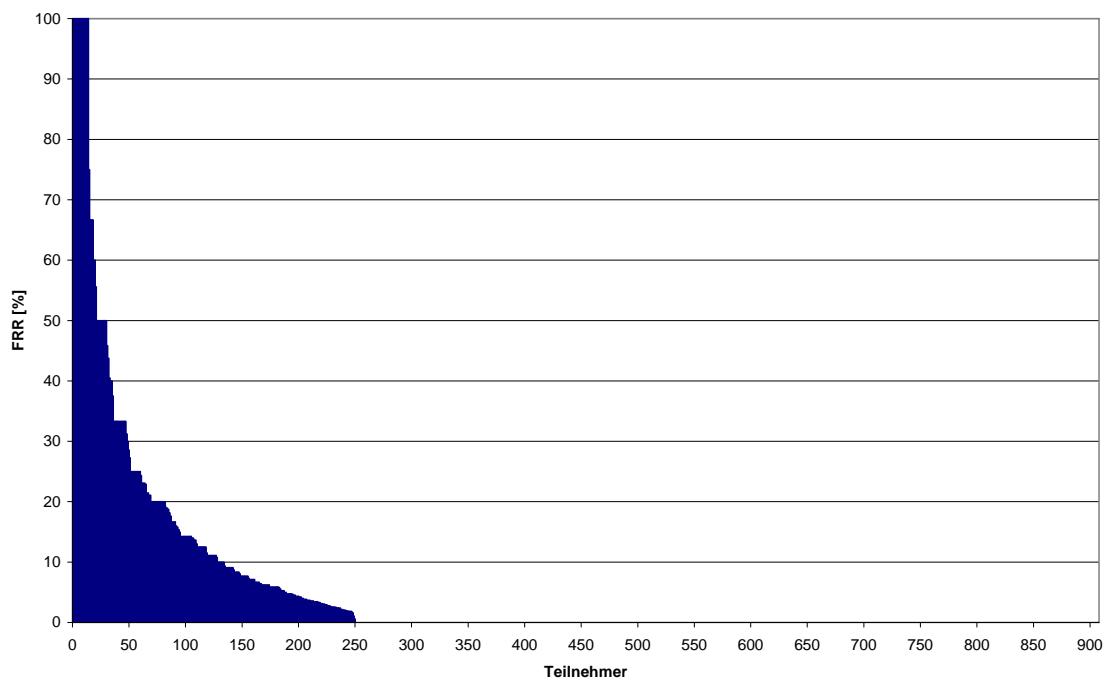


Abbildung 79: Einzelbenutzerstatistik RefID2 (Cognitec, Template, Zyklus 1), FAR=0,1%

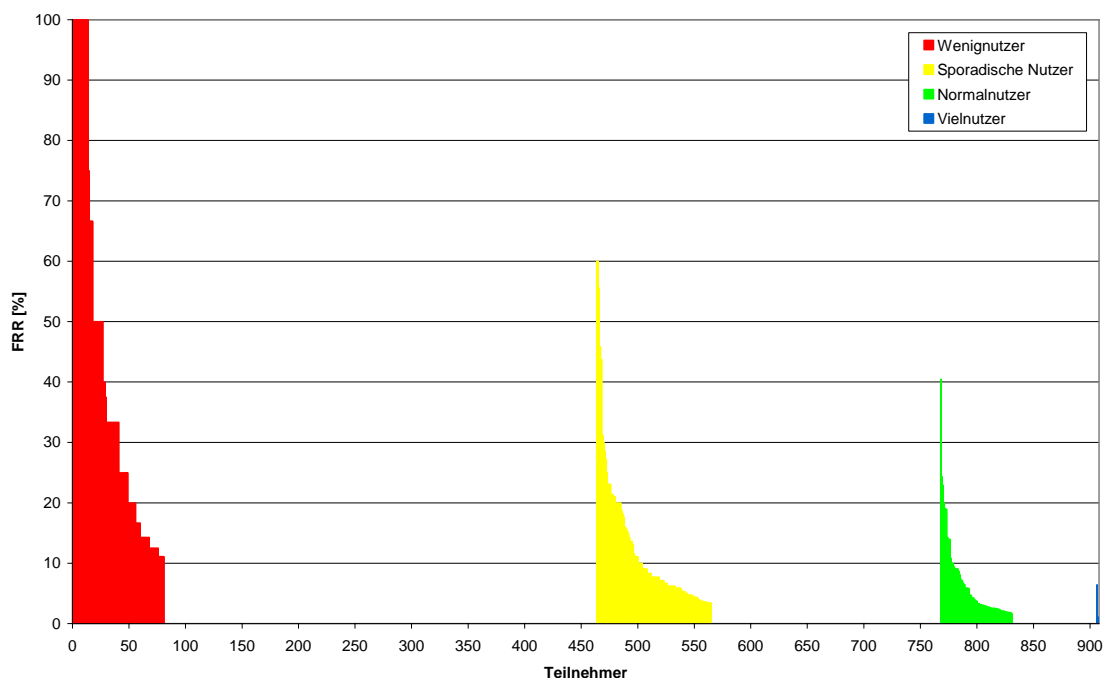


Abbildung 80: Nutzerklassenspezifische Einzelbenutzerstatistik RefID2 (Cognitec, Template, Zyklus 1), FAR=0,1%

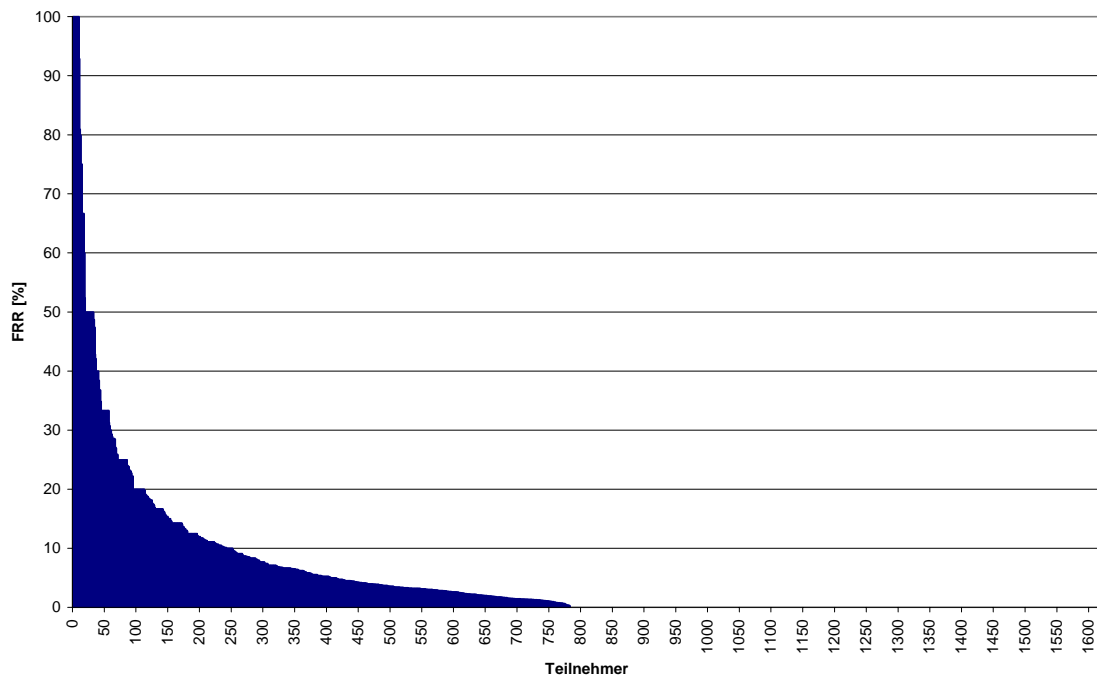


Abbildung 81: Einzelbenutzerstatistik RefID3 (BDr/NEC, ICAO-Bilder), FAR=0,081%

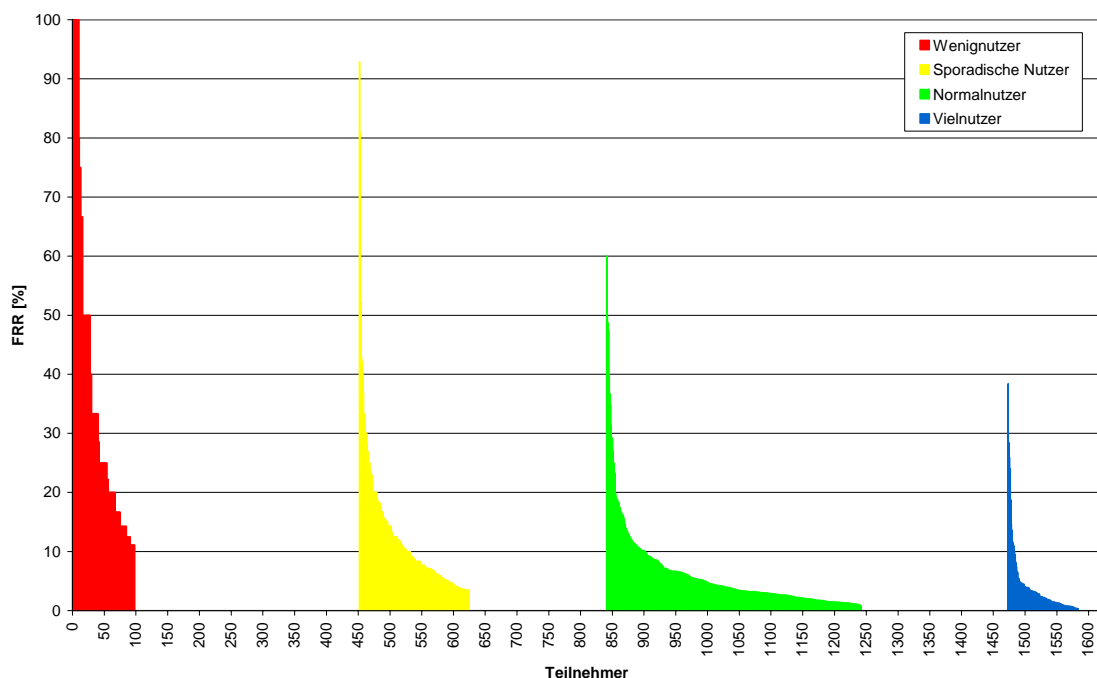


Abbildung 82: Nutzerklassenspezifische Einzelbenutzerstatistik RefID3 (BDr/NEC, ICAO-Bilder), FAR=0,081%

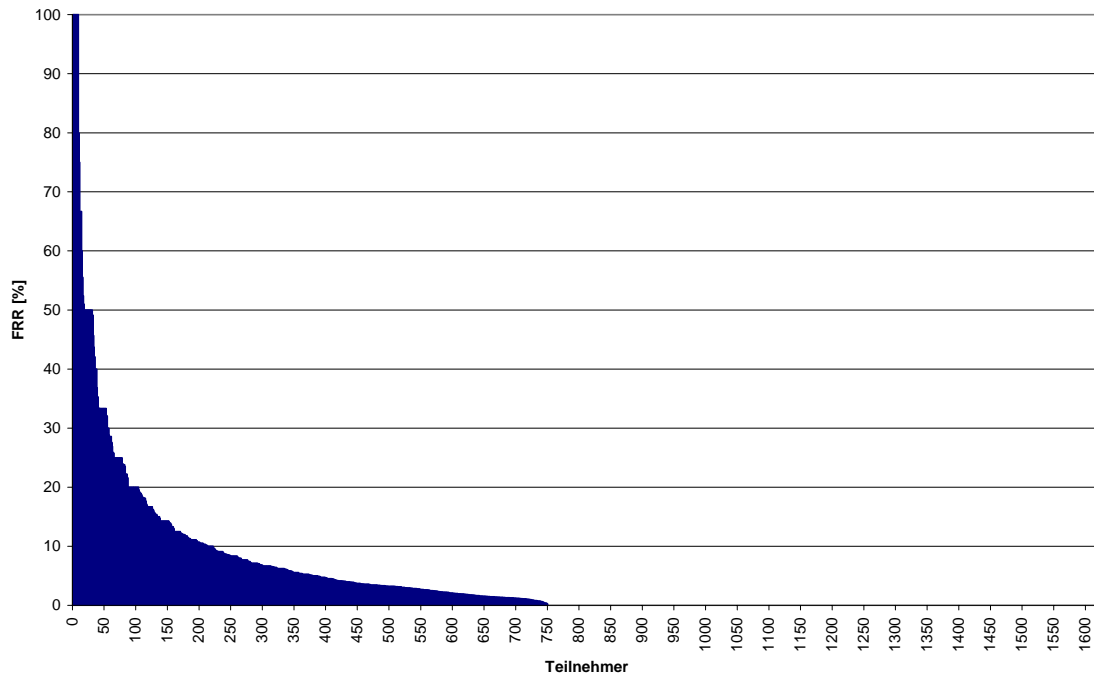


Abbildung 83: Einzelbenutzerstatistik RefID4 (BDr/NEC, Template), FAR=0,1%

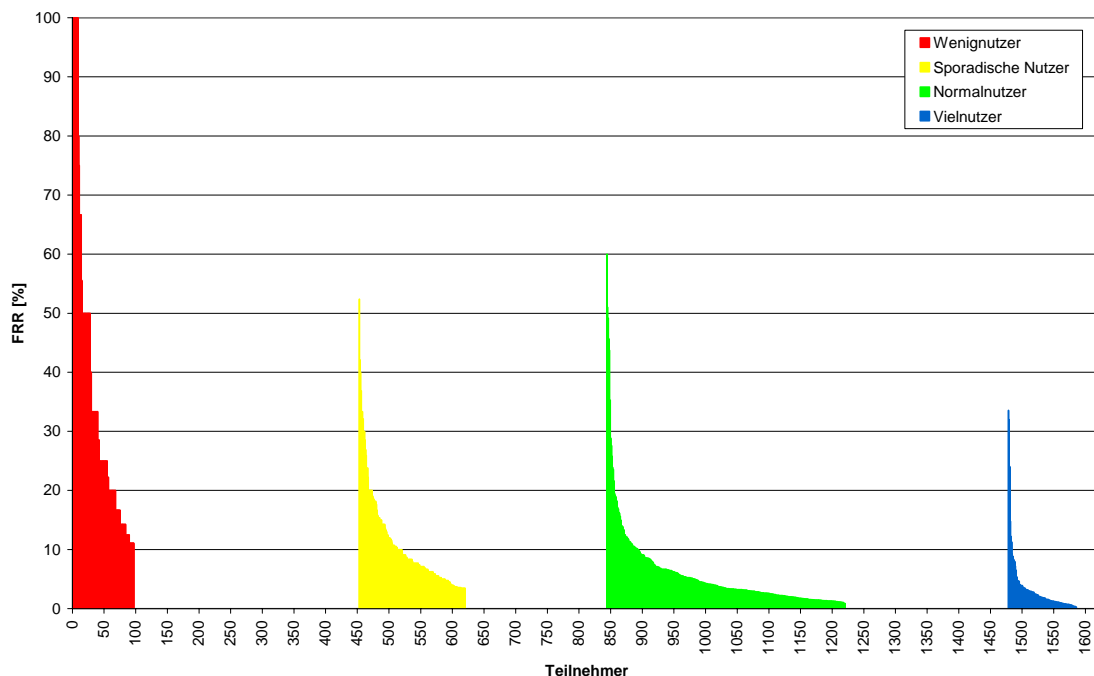


Abbildung 84: Nutzerklassenspezifische Einzelbenutzerstatistik RefID4 (BDr/NEC, Template), FAR=0,1%

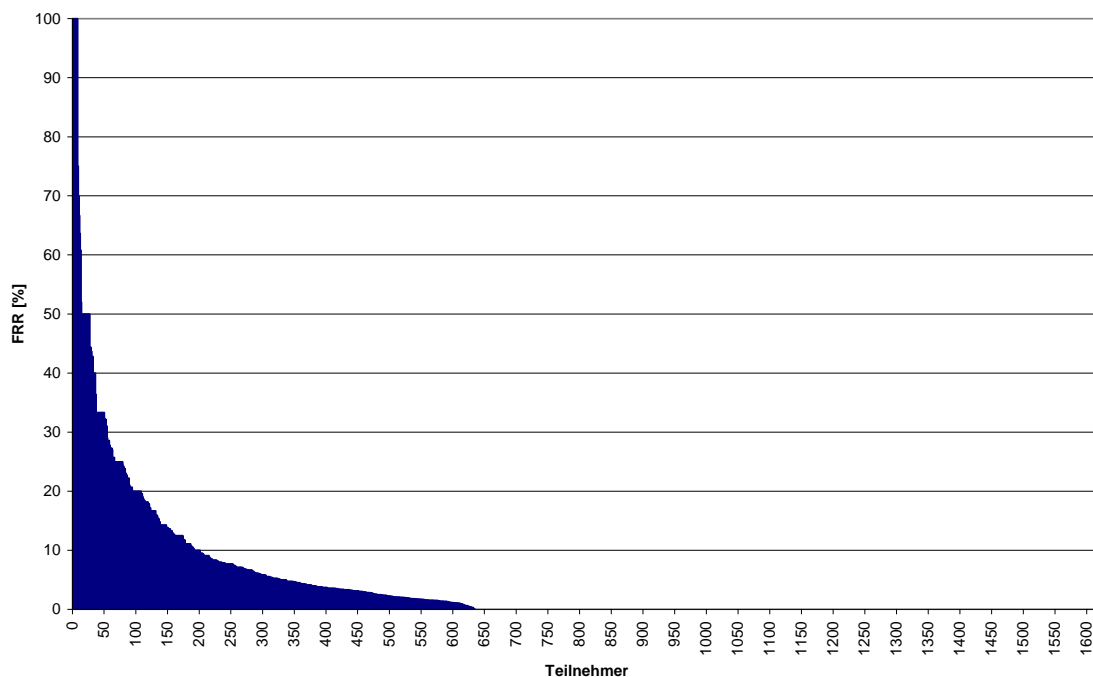


Abbildung 85: Einzelbenutzerstatistik RefID5 (Dermalog, ICAO-Bilder), FAR=0,1%

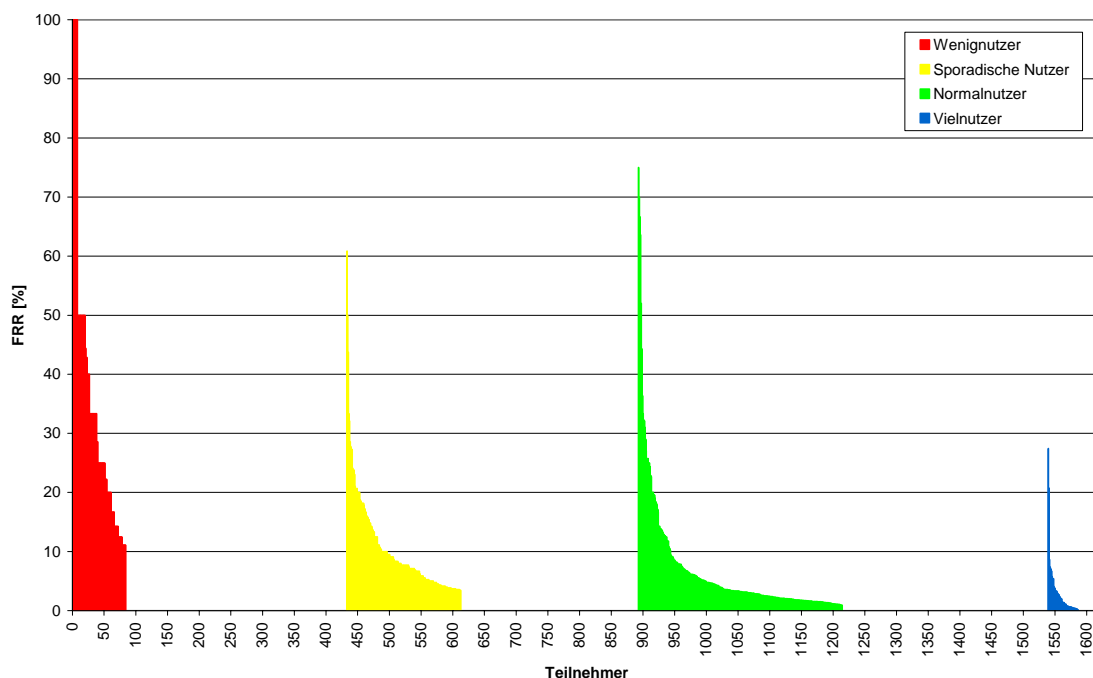


Abbildung 86: Nutzerklassenspezifische Einzelbenutzerstatistik RefID5 (Dermalog, ICAO-Bilder), FAR=0,1%

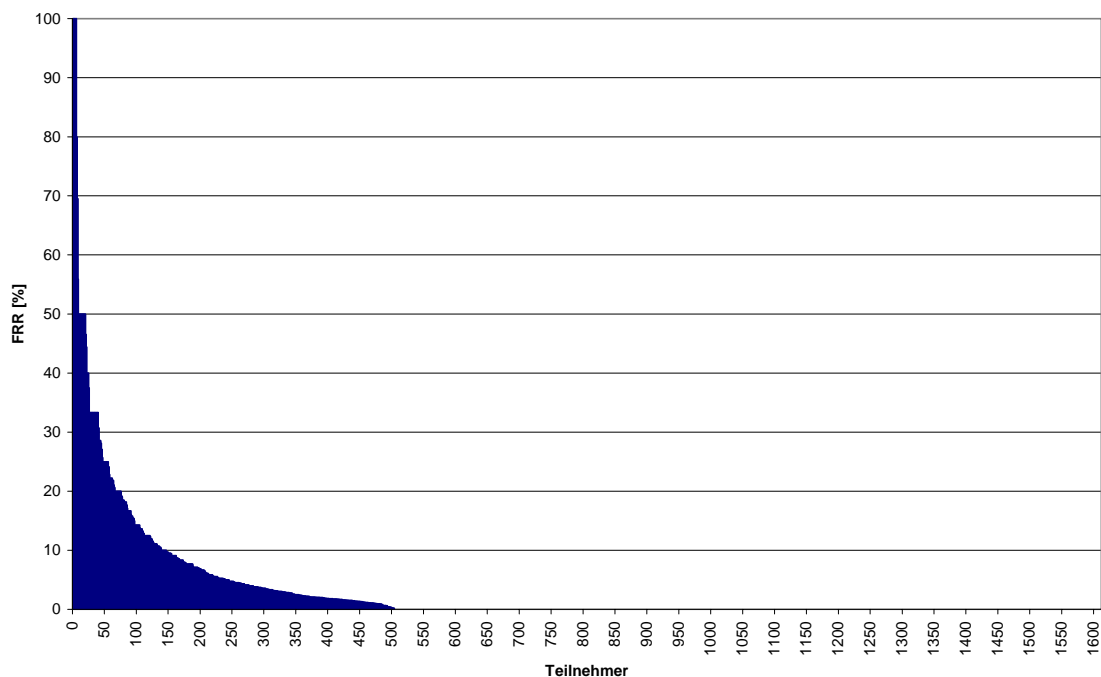


Abbildung 87: Einzelbenutzerstatistik RefID6 (Dermalog, Template), FAR=0,1%

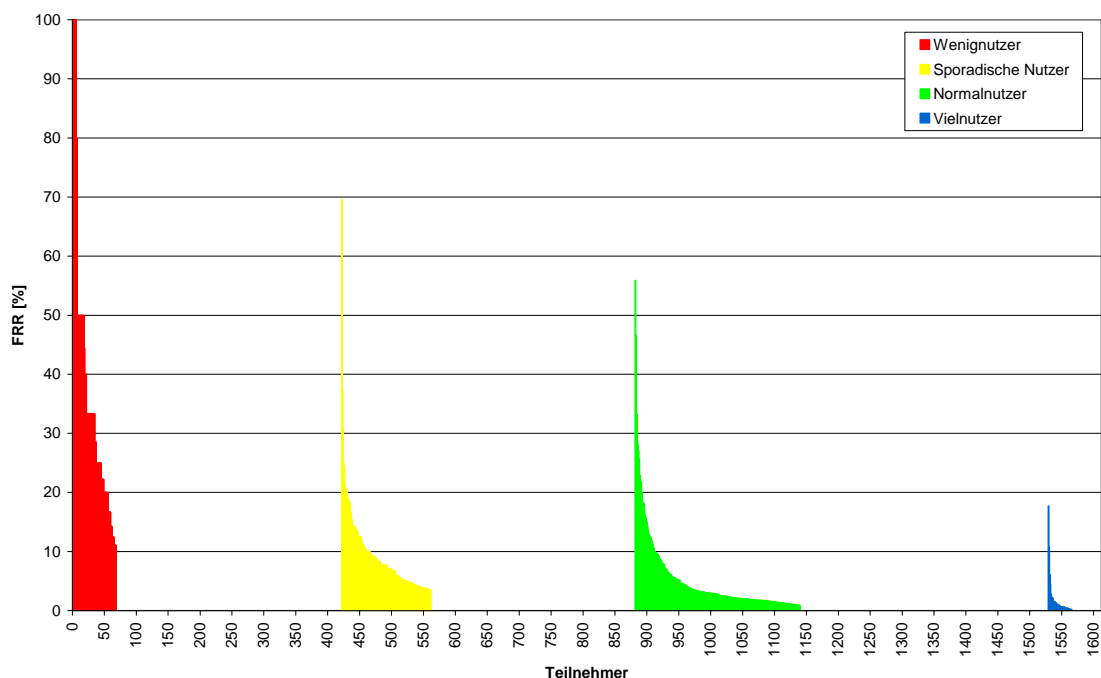


Abbildung 88: Nutzerklassenspezifische Einzelbenutzerstatistik RefID6 (Dermalog, Template), FAR=0,1%

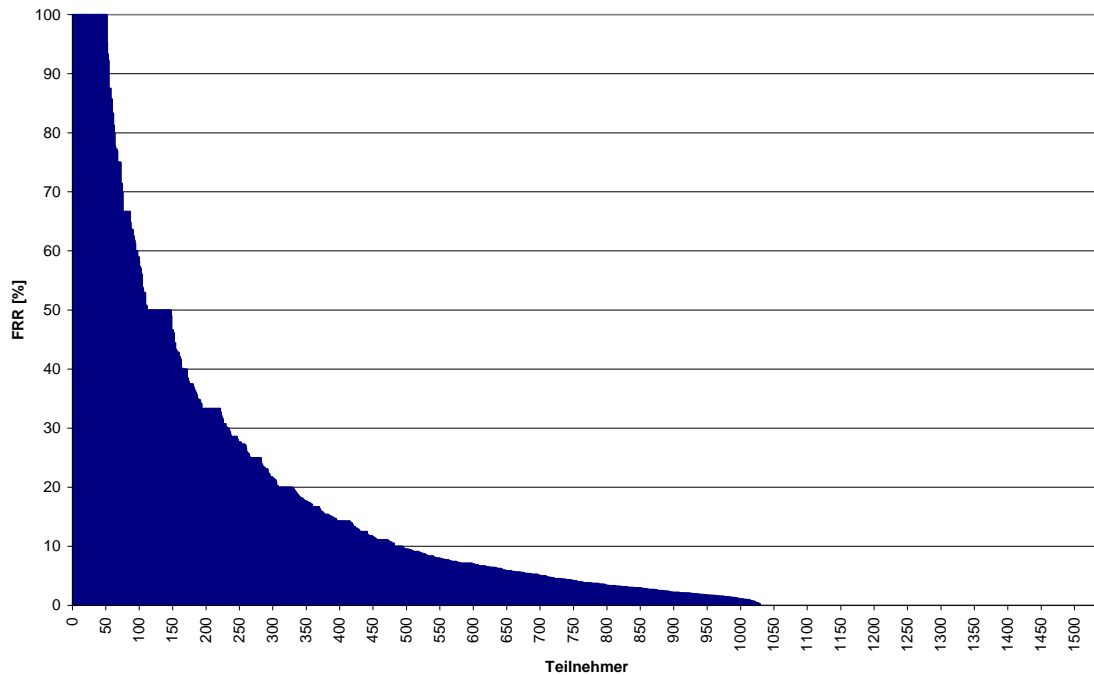


Abbildung 89: Einzelbenutzerstatistik RefID7 (SD Industries, ICAO-Bilder), FAR=0,1%

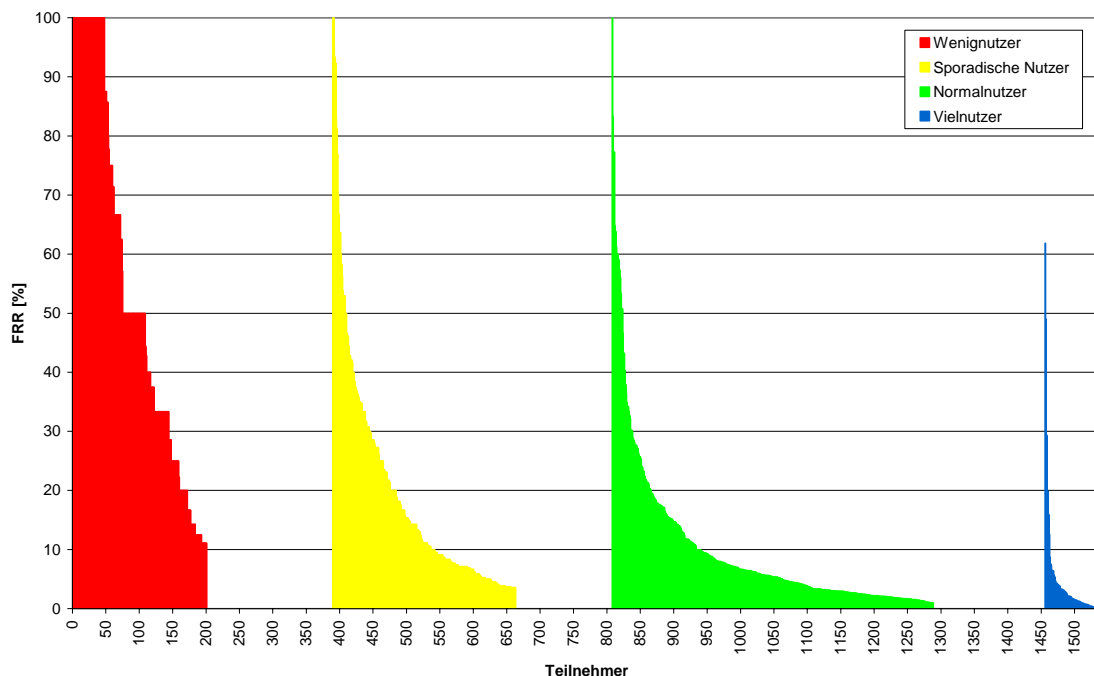


Abbildung 90: Nutzerklassenspezifische Einzelbenutzerstatistik RefID7 (SD Industries, ICAO-Bilder), FAR=0,1%

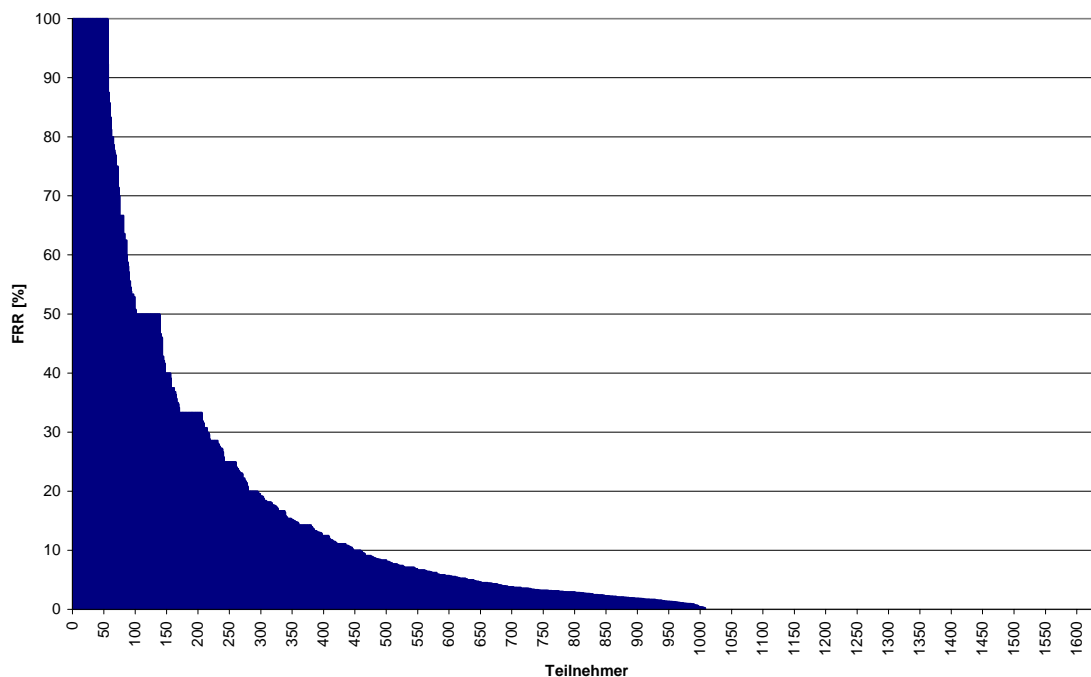


Abbildung 91: Einzelbenutzerstatistik RefID8 (SD Industries, Template), FAR=0,1%

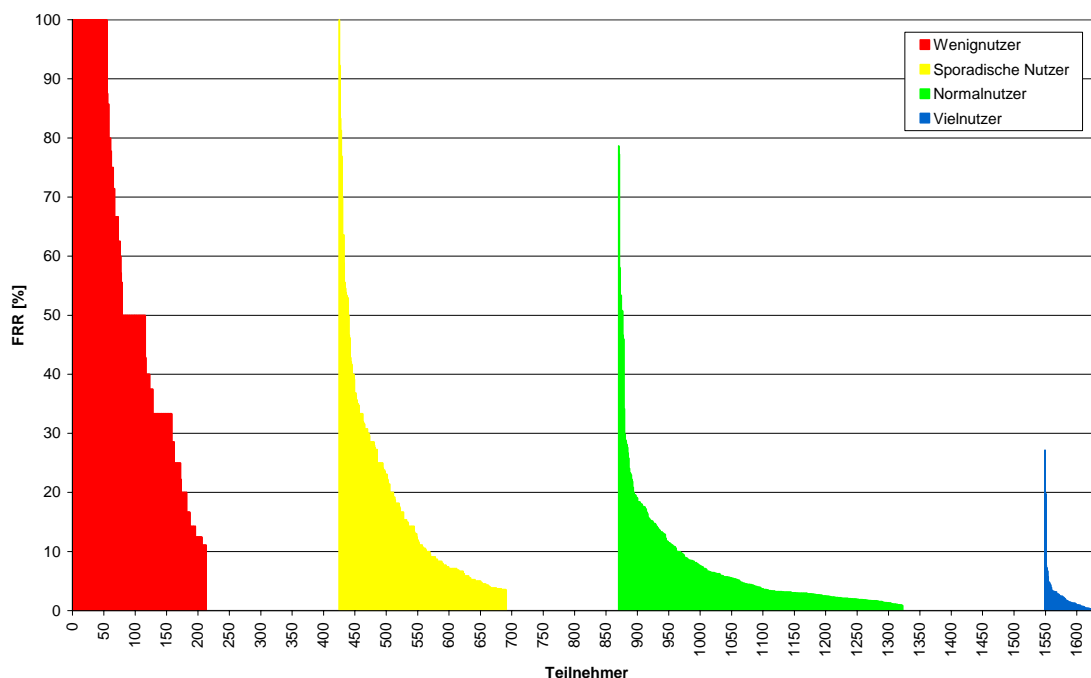


Abbildung 92: Nutzerklassenspezifische Einzelbenutzerstatistik RefID8 (SD Industries, Template), FAR=0,1%

Für SD Industries zeigt sich, dass ein großer Teil der Testteilnehmer erhebliche Probleme hatte. Bei beiden Referenzen gibt es über 450 Teilnehmer mit einer FRR von mindestens 10%, ein erheblicher Anteil erreicht sogar über 50%. Bei den Fingerabdruckererkennungssystemen zeigt sich ein deutlich anderes Bild. Während hier vergleichsweise wenige Teilnehmer extrem hohe FRRs aufweisen, ist für den überwiegenden Teil der Testpopulation die FRR gleich 0%.

Bei der Analyse der nutzerklassenspezifischen Einzelbenutzerstatistiken wird deutlich, dass der überwiegende Teil der Testteilnehmer mit sehr hohen FRRs in die Klassen der Wenignutzer und der sporadischen Nutzer fällt. Trotzdem gibt es auch bei den Normal- und Vielnutzern einzelne Personen, die hohe Rückweisungsraten aufweisen.

#### 7.2.3.4 Mittlere Einzelbenutzerstatistik (EBS)

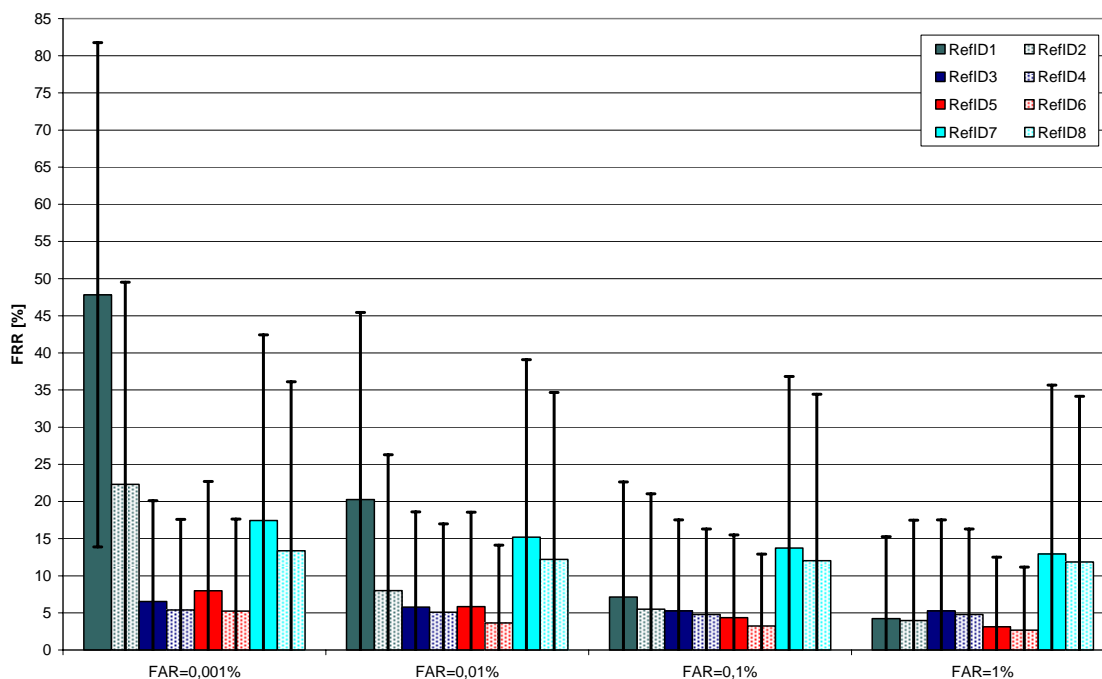
Um aus der Vielzahl der Einzelbenutzer-FRRs eine Kennzahl für jede RefID zu erhalten, wird jeweils das arithmetische Mittel der Einzelbenutzerstatistiken gebildet. Neben dem Mittelwert ist die Standardabweichung eine wichtige Kennzahl, da diese angibt, wie stark die einzelnen Werte um den Mittelwert schwanken.

Die Ergebnisse sind in Tabelle 28 und Abbildung 93 dargestellt.

RefID	FAR=0,001%		FAR=0,01%		FAR=0,1%		FAR=1%	
	EBS-Mittel	Standard-abw.	EBS-Mittel	Standard-abw.	EBS-Mittel	Standard-abw.	EBS-Mittel	Standard-abw.
1	47,82	33,95	20,25	25,20	7,13	15,50	4,23	11,04
2	22,29	27,22	7,99	18,30	5,51	15,52	3,95	13,52
3	6,54	13,56	5,79	12,81	5,28	12,26	5,28	12,26
4	5,39	12,20	5,12	11,86	4,79	11,51	4,78	11,51
5	7,98	14,72	5,85	12,71	4,36	11,12	3,12	9,39
6	5,25	12,37	3,67	10,46	3,22	9,70	2,66	8,51
7	17,44	25,00	15,19	23,91	13,73	23,09	12,95	22,72
8	13,36	22,75	12,21	22,47	12,04	22,38	11,84	22,32

**Tabelle 28: Mittlere FRR der Einzelbenutzerstatistik und Standardabweichung in [%]<sup>35</sup>**

<sup>35</sup> Bei RefID3 wird anstelle der Arbeitspunkte FAR=0,1% und FAR=1% der Arbeitspunkt FAR=0,081% betrachtet, bei RefID4 anstelle des Arbeitspunktes FAR=1% der Arbeitspunkt FAR=0,106%.



**Abbildung 93: Mittlere FRR der Einzelbenutzerstatistik und Standardabweichung<sup>36</sup>**

Es fällt auf, dass bei hohen Sicherheitsniveaus die Gesichtserkennung eine sehr hohe mittlere Falschrückweisungsrate mit großer Abweichung liefert. Die Iriserkennung hat für alle betrachteten Arbeitspunkte eine hohe mittlere Falschrückweisungsrate mit großer Standardabweichung. Unter Berücksichtigung der Tatsache, dass für Bundesdruckerei/NEC beim Arbeitspunkt FAR=1% Werte von deutlich stärkeren Sicherheitsniveaus zugrunde liegen (FAR=0,081% für RefID3 und FAR=0,106% für RefID4), bietet die Fingerabdruckerkennung für alle Arbeitspunkte das beste Verhalten.

### 7.2.3.5 Zusammenfassung Erkennungsleistung

Die Auswertungstypen Gesamterkennungsleistung, nutzerklassenspezifische Erkennungsleistung und mittlere Einzelbenutzerstatistik werden nachfolgend für die vier betrachteten Arbeitspunkte zusammenfassend dargestellt.

<sup>36</sup> Bei RefID3 wird anstelle der Arbeitspunkte FAR=0,1% und FAR=1% der Arbeitspunkt FAR=0,081% betrachtet, bei RefID4 anstelle des Arbeitspunktes FAR=1% der Arbeitspunkt FAR=0,106%.

Angaben in [%]	RefID1	RefID2	RefID3	RefID4	RefID5	RefID6	RefID7	RefID8
<b>FAR=0,001%</b>								
<b>Wenignutzer</b>	46,01	21,75	7,06	6,79	10,02	7,25	26,22	23,42
<b>Sporadische Nutzer</b>	48,49	21,11	6,72	5,26	8,29	4,77	15,27	10,92
<b>Normalnutzer</b>	49,07	22,16	5,66	4,11	6,85	4,24	11,62	6,24
<b>Vielnutzer</b>	45,05	15,20	5,09	3,09	4,50	1,69	7,80	3,27
<b>Mittlere EBS</b>	47,82	22,29	6,54	5,39	7,98	5,25	17,45	13,36
<b>Gesamt</b>	<b>47,13</b>	<b>21,03</b>	<b>5,55</b>	<b>3,86</b>	<b>6,40</b>	<b>3,62</b>	<b>11,27</b>	<b>6,36</b>
<b>FAR=0,01%</b>								
<b>Wenignutzer</b>	20,44	8,24	6,64	6,68	7,57	5,52	24,61	22,42
<b>Sporadische Nutzer</b>	22,21	6,44	6,05	4,98	5,96	3,41	13,39	9,88
<b>Normalnutzer</b>	20,14	6,19	4,62	3,69	4,74	2,43	8,89	4,91
<b>Vielnutzer</b>	17,25	3,16	3,85	2,69	2,82	0,81	5,60	2,04
<b>Mittlere EBS</b>	20,25	7,99	5,79	5,11	5,85	3,67	15,19	12,21
<b>Gesamt</b>	<b>18,91</b>	<b>6,02</b>	<b>4,48</b>	<b>3,47</b>	<b>4,39</b>	<b>2,15</b>	<b>8,83</b>	<b>5,10</b>
<b>FAR=0,1%<sup>37</sup></b>								
<b>Wenignutzer</b>	10,27	5,60	6,48	6,58	5,88	5,19	23,41	22,20
<b>Sporadische Nutzer</b>	8,57	4,20	5,52	4,59	4,18	2,95	11,94	9,73
<b>Normalnutzer</b>	6,10	3,58	3,91	3,29	3,20	1,95	7,26	4,75
<b>Vielnutzer</b>	4,77	1,82	2,90	2,27	1,65	0,66	4,20	1,97
<b>Mittlere EBS</b>	7,13	5,51	5,28	4,79	4,36	3,22	13,73	12,04
<b>Gesamt</b>	<b>5,67</b>	<b>3,66</b>	<b>3,70</b>	<b>3,07</b>	<b>2,93</b>	<b>1,78</b>	<b>7,30</b>	<b>4,97</b>
<b>FAR=1%<sup>38</sup></b>								
<b>Wenignutzer</b>	6,78	3,59	6,48	6,58	4,41	4,52	22,72	22,10
<b>Sporadische Nutzer</b>	4,90	3,04	5,52	4,59	2,90	2,58	10,97	9,47
<b>Normalnutzer</b>	3,66	2,12	3,91	3,29	1,96	1,46	6,38	4,51
<b>Vielnutzer</b>	3,01	1,03	2,90	2,25	0,90	0,48	3,28	1,86
<b>Mittlere EBS</b>	4,23	3,95	5,28	4,78	3,12	2,66	12,95	11,84
<b>Gesamt</b>	<b>3,47</b>	<b>2,33</b>	<b>3,70</b>	<b>3,06</b>	<b>1,83</b>	<b>1,39</b>	<b>6,40</b>	<b>4,78</b>

**Tabelle 29: FRR bei unterschiedlichen Sicherheitsniveaus**

<sup>37</sup> RefID3 bei Arbeitspunkt FAR=0,081%

<sup>38</sup> RefID3 bei Arbeitspunkt FAR=0,081% und RefID4 bei Arbeitspunkt FAR=0,106%

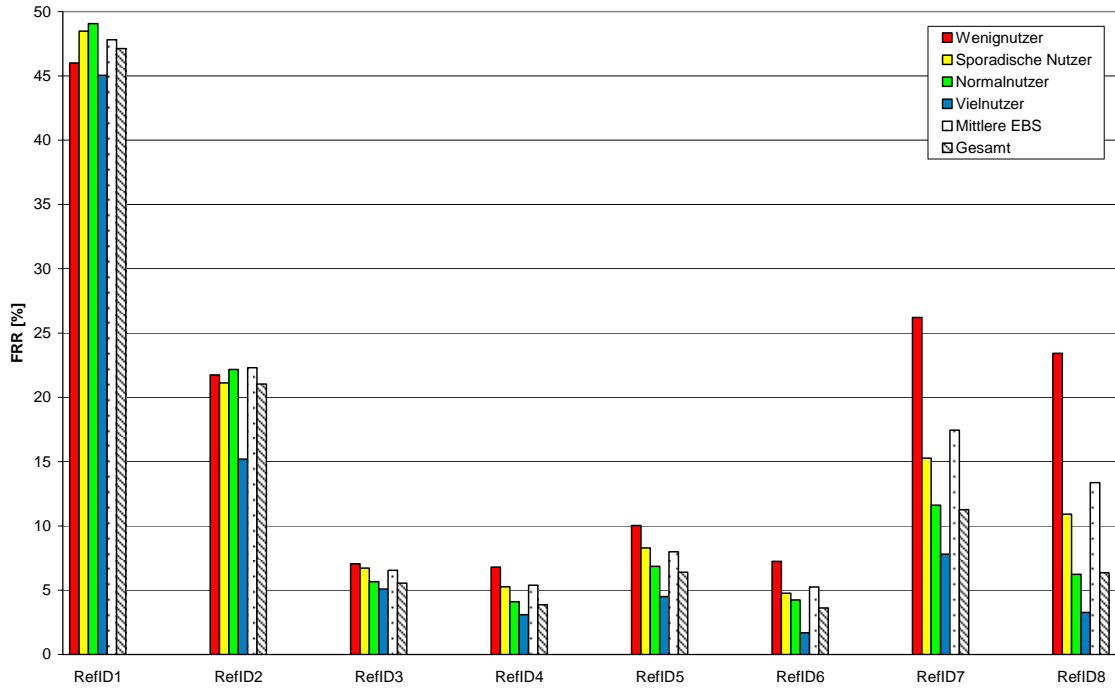


Abbildung 94: FRR bei Arbeitspunkt FAR=0,001%

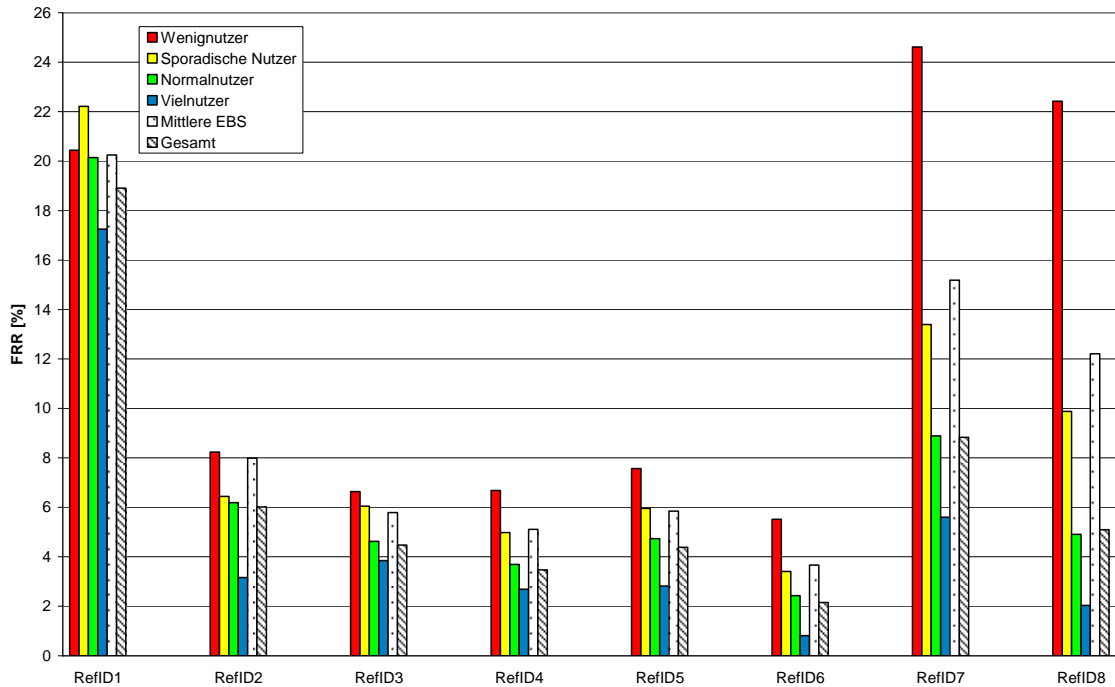


Abbildung 95: FRR bei Arbeitspunkt FAR=0,01%

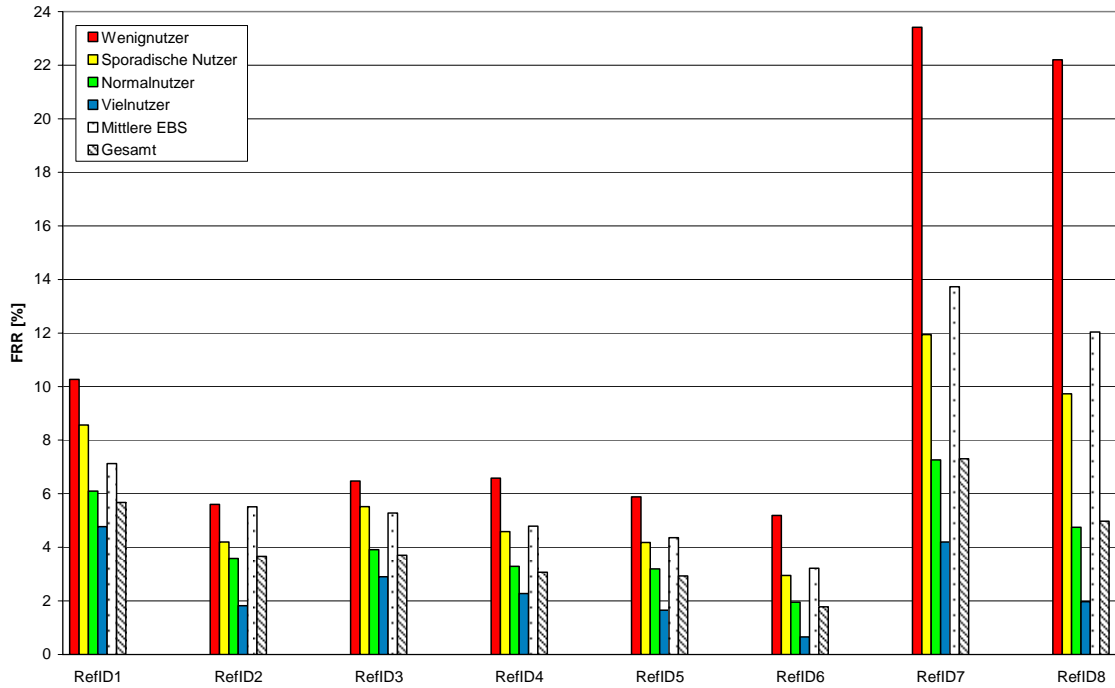


Abbildung 96: FRR bei Arbeitspunkt FAR=0,1%<sup>39</sup>

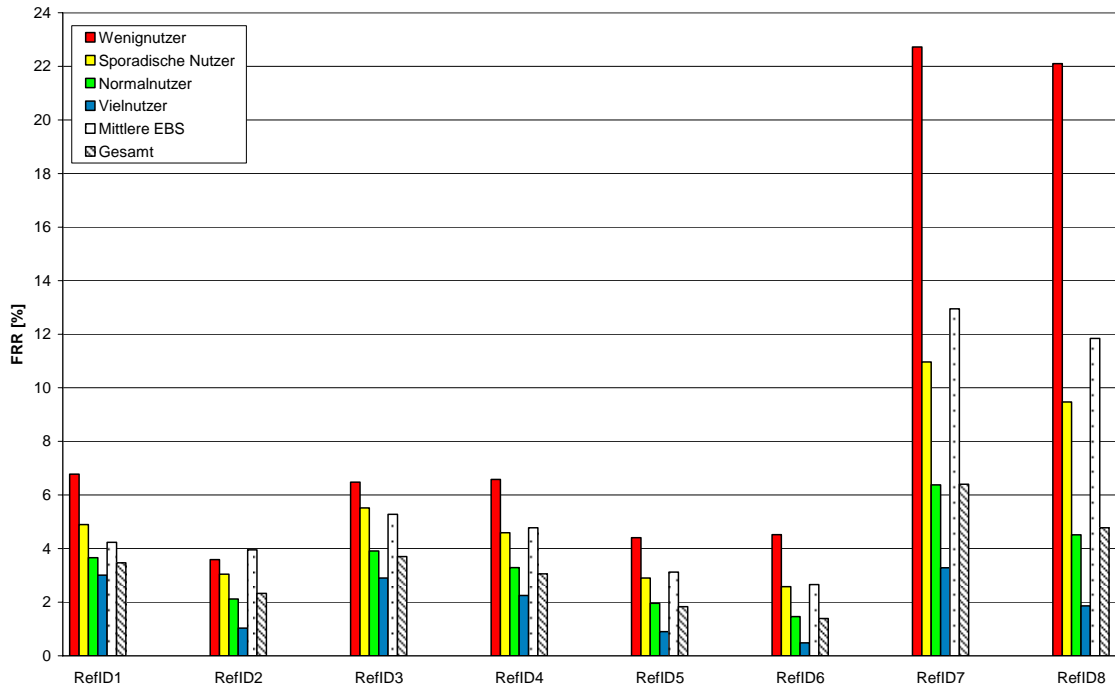


Abbildung 97: FRR bei Arbeitspunkt FAR=1%<sup>40</sup>

<sup>39</sup> RefID3 bei Arbeitspunkt FAR=0,081%

Für alle Sicherheitsniveaus wird deutlich, dass die Gesamt-FRR näherungsweise der FRR der Gruppe der Normalnutzer entspricht. Die mittlere FRR der Einzelbenutzerstatistik ist zum Teil deutlich größer als die Gesamt-FRR. Dies rührt daher, dass Wenignutzer, die nur eine oder mehr Betätigungen durchgeführt haben und in der Regel dabei sehr hohe Rückweisungsrate erzielten, mit dem selben Gewicht eingehen wie Teilnehmer, die regelmäßig Verifikationen ohne Rückweisungen hatten.

In den folgenden Abbildungen ist die mittlere Einzelbenutzerstatistik in die DET-Darstellung der Gesamterkennungsleistung eingeordnet (Abbildung 98 entspricht somit Abbildung 57 mit zusätzlichen diskreten Werten an den vier betrachteten Arbeitspunkten; Abbildung 99 entspricht Abbildung 58 inkl. mittlerer EBS).

Dabei ist zu berücksichtigen, dass für Bundesdruckerei/NEC die am Arbeitspunkt FAR=1% dargestellten FRR-Werte der mittleren Einzelbenutzerstatistik ein deutlich höheres Sicherheitsniveau repräsentieren (FAR=0,081% für RefID3 und FAR=0,106% für RefID4).

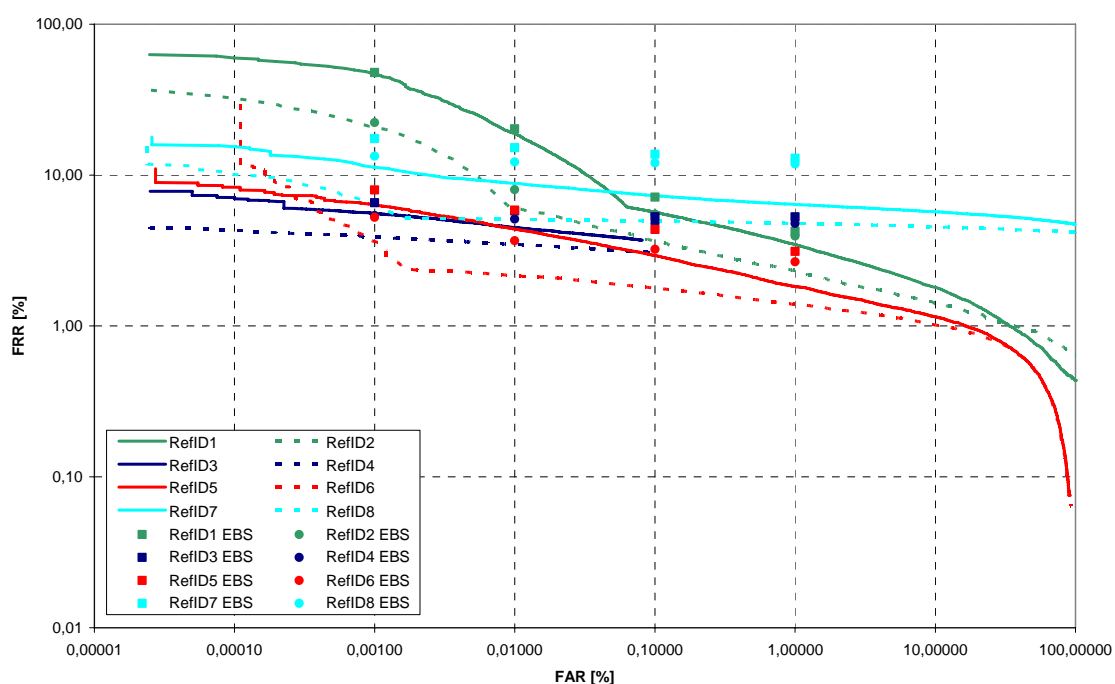
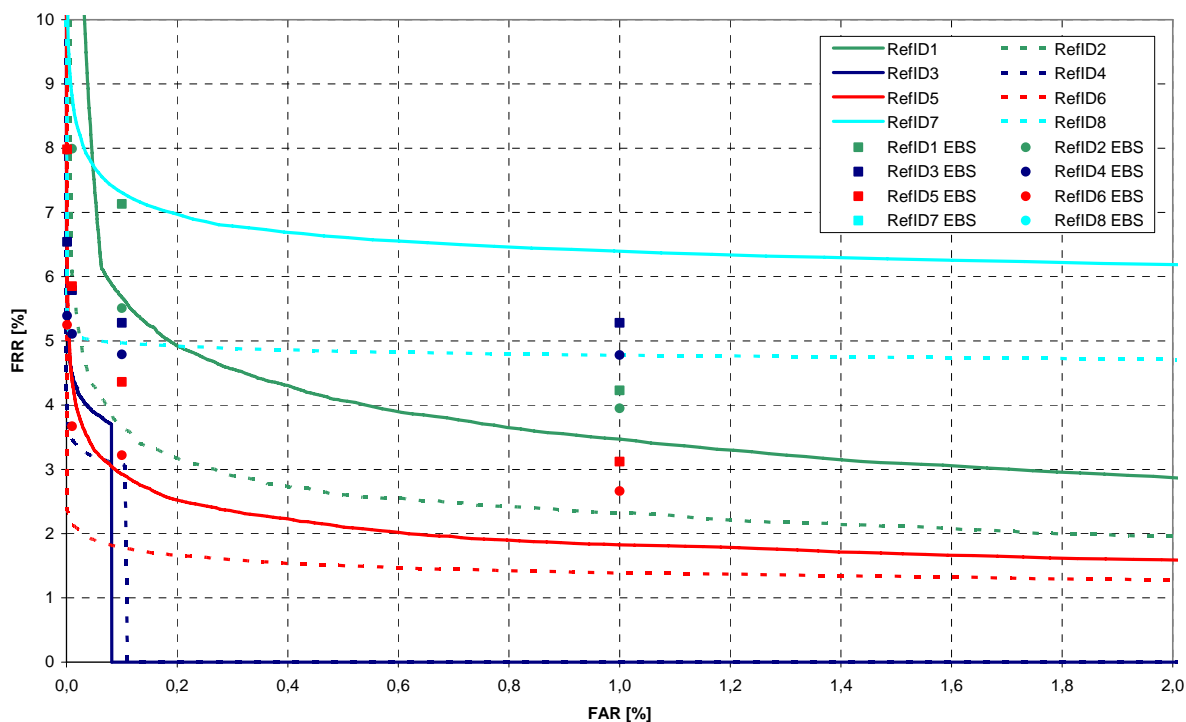


Abbildung 98: DET-Kurven inkl. mittlerer Einzelbenutzerstatistik (logarithmische Skalierung des Arbeitsbereichs)

<sup>40</sup> RefID3 bei Arbeitspunkt FAR=0,081% und RefID4 bei Arbeitspunkt FAR=0,106%



**Abbildung 99: DET-Kurven über den Arbeitsbereich FAR ≤ 2%, FRR ≤ 10% inkl. mittlerer Einzelbenutzerstatistik**

#### 7.2.4 Generalisierte Erkennungsleistungen

Bei jeder Betätigung wird aus dem aufgenommenen Live-Bild ein Template gebildet. Schlägt diese Template-Bildung fehl, wird dies in einer realen Einsatzumgebung als Falschrückweisung interpretiert. Der Anteil von Personen dar, bei denen aus dem biometrischen Merkmal kein Template gebildet kann, wird durch die FTE repräsentiert. Daher ist die FTE bei der Bewertung der Erkennungsleistung mit zu berücksichtigen.

Eine Möglichkeit der zusammenfassenden Darstellung von FTE und FRR bzw. FAR ist gemäß [ISO\_BPTR] die Verwendung von generalisierten Falschrückweisungs- bzw. Falschakzeptanzraten. In den generalisierten Fehlerraten werden die Fehler bei Merkmalsaufnahme (Failure-To-Acquire - FTA), Enrolment (Failure-To-Enrol - FTE) und Vergleich (False Match Rate - FMR bzw. False Non Match Rate - FNMR) zusammengefasst.

##### 7.2.4.1 Generalisierte FAR

Für eine fälschliche Akzeptanz eines Unberechtigten müssen folgende Bedingungen erfüllt sein:

1. Ein Berechtigter muss enroled sein.
2. Das Merkmal des Unberechtigten muss erfasst werden können.
3. Aus dem Merkmal des Unberechtigten muss ein Template generiert werden können (Enrolment).
4. Es muss ein erfolgreicher Vergleich des Templates des Berechtigten mit dem Template des Unberechtigten erfolgen.

Da alle Bedingungen gemeinsam erfüllt sein müssen, berechnet sich die generalisierte FAR wie folgt:

$$FAR_{Gen} = (1 - FTE) (1 - FTA) (1 - FTE) FMR$$

$$FAR_{Gen} = FMR (1 - FTA) (1 - FTE)^2$$

$$FAR_{Gen} = FAR (1 - FTE)^2$$

Tabelle 30 stellt die generalisierte FAR für ein Sicherheitsniveau von FAR=0,1% dar.

System	RefID	FTE [%]	FAR <sub>Gen</sub> [%]
Cognitec (GE)	1 (ICAO-Bild)	0	0,1
	2 (Template)	0	0,1
BDr/NEC (FE)	3 (ICAO-Bild)	0,25	0,0995
	4 (Template)	0,99	0,0980
Dermalog (FE)	5 (ICAO-Bild)	0	0,1
	6 (Template)	0,40	0,0992
SDI (IE)	7 (ICAO-Bild)	0	0,1
	8 (Template)	0,89	0,0982

**Tabelle 30: Generalisierte FAR<sub>Gen</sub> bei FAR=0,1%**

#### 7.2.4.2 Generalisierte FRR

Eine fälschliche Abweisung eines Berechtigten kann aus folgenden Gründen erfolgen:

1. Der Berechtigte kann nicht enroled werden.
2. Das Merkmal des Berechtigten kann bei der Verifikation nicht erfasst werden.
3. Der Vergleich des Referenz-Templates mit dem Template aus der Live-Aufnahme schlägt fehl.

Da für eine falsche Rückweisung nur einer der angegebenen Punkte erfüllt sein muss, berechnet sich die generalisierte FRR wie folgt:

$$FRR_{Gen} = FTE + (1 - FTE) FTA + (1 - FTE) (1 - FTA) FNMR$$

$$FRR_{Gen} = FTE + (1 - FTE) (FTA + (1 - FTA) FNMR)$$

$$FRR_{Gen} = FTE + (1 - FTE) FRR$$

Tabelle 31 stellt die generalisierte FRR für ein Sicherheitsniveau von FAR=0,1% dar.

System	RefID	FTE [%]	FRR [%]	FRR <sub>Gen</sub> [%]
Cognitec (GE)	1 (ICAO-Bild)	0	5,67	5,67
	2 (Template)	0	3,66	3,66
BDr/NEC (FE)	3 (ICAO-Bild)	0,25	3,70	3,94
	4 (Template)	0,99	3,07	4,03
Dermalog (FE)	5 (ICAO-Bild)	0	2,93	2,93
	6 (Template)	0,40	1,78	2,17
SDI (IE)	7 (ICAO-Bild)	0	7,30	7,30
	8 (Template)	0,89	4,97	5,82

**Tabelle 31: Generalisierte FRR<sub>Gen</sub> bei FAR=0,1%<sup>41</sup>**

### 7.2.5 Korrelationen

Einen Bestandteil der Untersuchung von biometrischen Systemen im Rahmen von BioP II bildet die Betrachtung des Zusammenhangs zwischen der Erkennungsleistung (FRR) und demographischen (Geschlecht, Alter etc.) bzw. merkmalsbeeinflussenden (Brille, Bart, Kontaktlinsen) Eigenschaften. Ziel der Betrachtung ist die Analyse von möglichen Zusammenhängen zwischen den genannten Zielgrößen sowie die Identifizierung etwaiger Ursachen. Die wesentlichen Resultate der Untersuchungen sind im Folgenden dargestellt.

Grundlage der Korrelationsuntersuchungen sind die Verifikationsergebnisse beim Arbeitspunkt FAR=0,1%.<sup>42</sup> Für die durchgeführten Untersuchungen gilt, dass das verwendete statistische Verfahren mit einer Wahrscheinlichkeit von 99% das korrekte Ergebnis liefert.

#### 7.2.5.1 Geschlecht

Zwischen dem Merkmal Geschlecht und den Falschrückweisungsrate besteht sowohl bei dem Gesichtserkennungssystem als auch bei dem Fingerabdruckererkennungssystem von Bundesdruckerei/NEC eine Abhängigkeit. In diesen Fällen weisen die weiblichen Testpersonen eindeutig mehr und die männlichen weniger Erkennungsfehler auf als bei Unabhängigkeit zu erwarten wären. Beim Fingerabdruckererkennungssystem der Firma Dermalog sowie dem Iriserkennungssystem konnten keine Abhängigkeiten zwischen dem Geschlecht und den Falschrückweisungsrate festgestellt werden.

Für Cognitec bestätigt dies die Beobachtung aus BioP I (siehe [BioPI]). Dort war ebenfalls als Tendenz zu erkennen, dass männliche Testteilnehmer geringere Rückweisungsrate aufwiesen als weibliche.

Bemerkenswert ist, dass nur für eines der beiden Fingerabdruckererkennungssysteme (BDr/NEC) die weiblichen Testteilnehmer signifikant höhere Rückweisungsrate aufweisen als die männlichen. Bei BDr/NEC wird das Live-Bild des Fingers auf eine Mindestanzahl von Minuten überprüft. Wird diese unterschritten, führt das zu einem negativen Verifikationsergebnis (Matchscore = 0). Da Frauen in der Regel kleinere Finger haben als Männer und somit die Anzahl der zu erfassenden Minuten geringer

<sup>41</sup> Für RefID 3 ist der Arbeitspunkt FAR=0,081%.

<sup>42</sup> Für RefID 3 ist der Arbeitspunkt FAR=0,081%.

ist, führt dies vermutlich zu der beobachteten schlechteren Erkennungsleistung bei den weiblichen Testteilnehmern.

#### *7.2.5.2 Alter*

Beim Gesichtserkennungssystem ist kein direkter Zusammenhang zwischen den Falschrückweisungs-raten und der Altersstruktur zu sehen. Es existieren einzelne Abhängigkeiten, die sich jedoch nicht eindeutig einem Altersbereich zuordnen lassen.

Bei den Fingerabdruckerkennungssystemen ist auffällig, dass durchweg eine Abhängigkeit besteht. Es gibt signifikant höhere Falschrückweisungs-raten der älteren Testpersonen, als unter der Unabhängigkeitsprämisse zu erwarten wären. Dagegen weist die jüngere Testpopulation durchweg geringere Erkennungsfehler auf, als bei Unabhängigkeit zu erwarten wären. Dies Ergebnis ist zu erwarten, da bei älteren Personen die Haut in der Regel trockener und damit das Bild des Fingerabdrucks weniger kontrastreich ist. Eine weitere mögliche Ursache ist, dass bei älteren Personen die Finger abgegriffener sind und somit die Strukturen schwächer ausgeprägt sind. Möglicherweise spielen auch Bedienungsschwierigkeiten eine Rolle.

Bei dem Iriserkennungssystem lässt sich nicht eindeutig bestimmen, aus welchem Bereich der Altersstruktur die Abhängigkeit hervorgeht. Tendenziell weisen aber die Altersgruppen über 50 Jahre und unter 20 Jahre schlechtere Erkennungsraten auf als bei Unabhängigkeit zu erwarten wäre, die 30 bis 50-Jährigen bessere. Für die über 50-Jährigen kann dieser Trend zum Teil dadurch erklärt werden, dass der Anteil der Brillenträger bei älteren Personen deutlich höher ist und Brillen zu einer Verschlechterung der Erkennungsleistung führen (siehe Abschnitt 7.2.5.7). Des Weiteren spielen hier vermutlich auch Bedienungsschwierigkeiten eine Rolle.

#### *7.2.5.3 Ethnische Herkunft*

Die Zusammensetzung der Testpersonen schließt es aus, eine signifikante Aussage zu treffen. Die außereuropäischen Testpersonen stellen einen sehr geringen Anteil am Testumfang dar (ca. 2%). Somit führen schon geringfügige Änderungen bei diesen Testpersonen zu einer Veränderung der Abhängigkeitsbeziehung. Deshalb können keine verlässlichen Ergebnisse aus den Untersuchungen gewonnen werden.

#### *7.2.5.4 Ausbildung*

Bezüglich des Merkmals Berufsausbildung existieren einzelne Abhängigkeiten. Diese sind jedoch so uneinheitlich, dass hieraus keine belastbaren Aussagen abgeleitet werden können.

#### *7.2.5.5 Berufliche Tätigkeit*

Beim Gesichtserkennungssystem ist ein interessanter Sachverhalt sichtbar. Während im Testzyklus 1 administrativ Tätige signifikant bessere Erkennungsleistungen zeigen als operativ Beschäftigte, ist im Testzyklus 2 ein umgekehrtes Verhältnis zu beobachten.

Bei dem Fingerabdruckerkennungssystem von Bundesdruckerei/NEC weisen operativ Tätige eine signifikant höhere Falschrückweisungsrate auf, als bei Unabhängigkeit zu erwarten wäre. Für das System von Dermalog kann hingegen keine eindeutige Abhängigkeit zwischen Tätigkeit und Erkennungsleistung nachgewiesen werden.<sup>43</sup> Vermutlich greift hier ebenfalls der Effekt, dass BDr/NEC eine Qualitätsbewertung des aufgenommenen Bildes hinsichtlich der Mindestanzahl Minuten vornimmt. Bei operativen Mitarbeitern sind durch ihre Tätigkeit die Strukturen der Finger

---

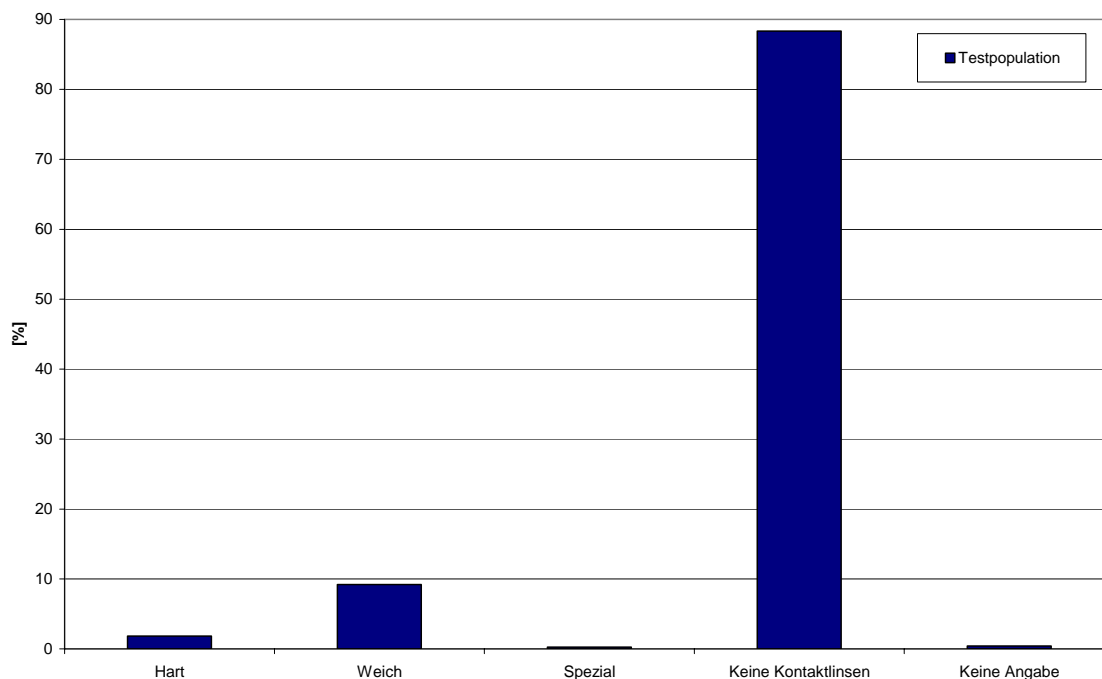
<sup>43</sup> Es existieren einzelne Abhängigkeiten, die sich jedoch nicht eindeutig zuordnen lassen.

möglicherweise schwächer ausgeprägt, so dass die minimal erforderliche Anzahl von Minuten nicht immer erreicht wird.

Beim Iriserkennungssystem erzielen die administrativ Beschäftigten signifikant bessere Erkennungsleistungen als bei Unabhängigkeit zu erwarten ist. Für die operativen Mitarbeiter gilt das Gegenteil, sie erreichen schlechtere Falschrückweisungsrate. Dies liegt vermutlich am Trainingseffekt. Da operativ Beschäftigte stärkeren zeitlichen Restriktionen unterliegen, sind diese eher in der Gruppe der sporadischen oder Normalnutzer vertreten während die Gruppe der Vielnutzer zum überwiegenden Teil aus administrativen Mitarbeitern besteht. Wie Abschnitt 7.2.3.2 zeigt, hat insbesondere bei der Iriserkennung die Gruppe der Vielnutzer eine erheblich bessere Erkennungsleistung als die anderen Nutzerklassen.

#### 7.2.5.6 Kontaktlinsen

Für die Iriserkennung wird untersucht, ob sich ein Zusammenhang zwischen dem Tragen von Kontaktlinsen und der Erkennungsleistung aufzeigen lässt. Dabei ist zu berücksichtigen, dass der überwiegende Teil der Testteilnehmer keine Kontaktlinse trägt (siehe Abbildung 100).

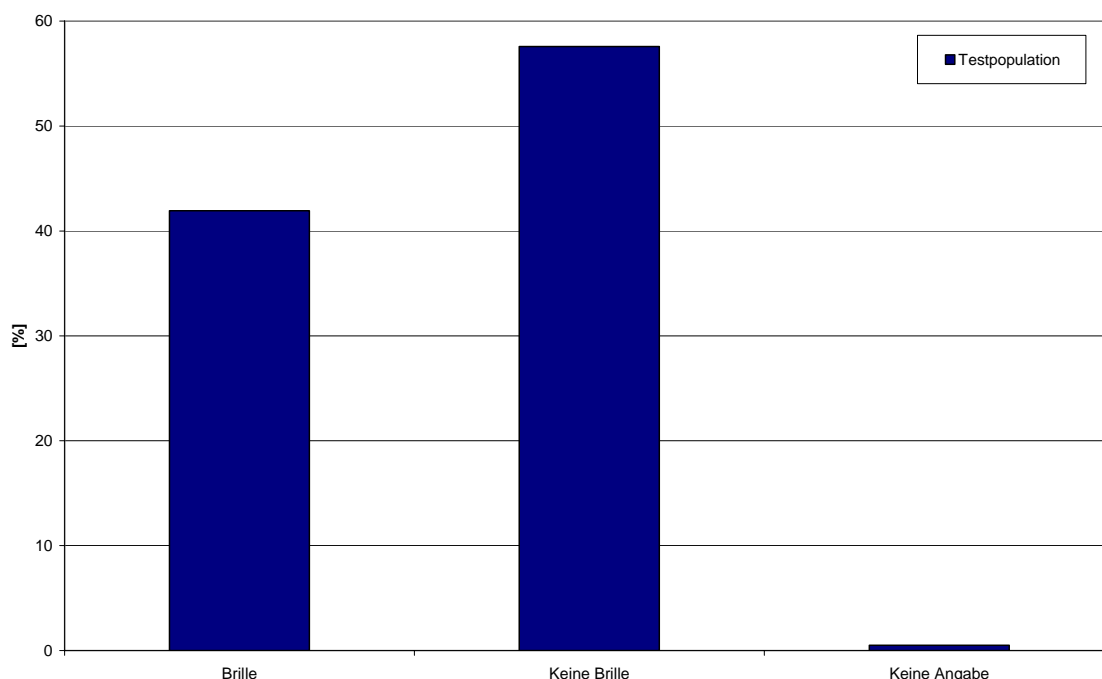


**Abbildung 100: Verteilung der Testpopulation hinsichtlich Merkmal Kontaktlinsen**

Bei den Korrelationsuntersuchungen zeigte sich ein uneinheitliches Verhalten. Eine Abhängigkeit zwischen der Falschrückweisungsrate und dem Vorhandensein von Kontaktlinsen kann weder bestätigt noch widerlegt werden.

#### 7.2.5.7 Brille

Sowohl für die Gesichts- als auch für die Iriserkennung wird untersucht, ob sich ein Zusammenhang zwischen dem Tragen von Brillen und der Erkennungsleistung aufzeigen lässt. Der Anteil der Brillenträger an der Testpopulation ist in Abbildung 101 dargestellt.



**Abbildung 101: Verteilung der Testpopulation hinsichtlich Merkmal Brille**

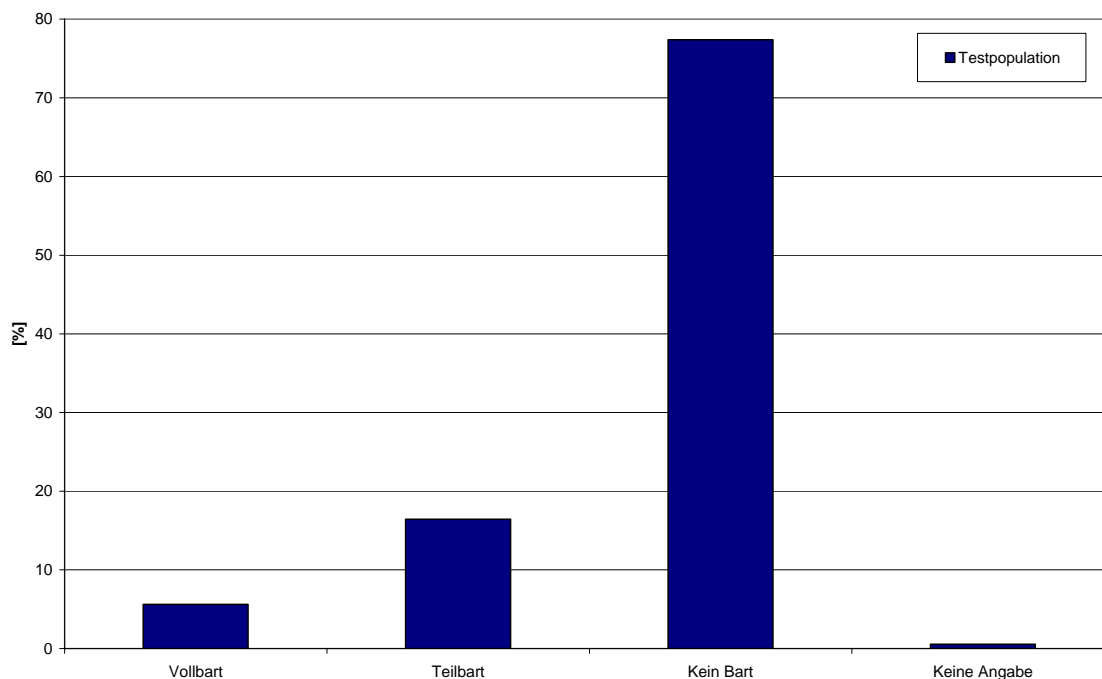
Für die Gesichtserkennung lässt sich festhalten, dass für das Merkmal Brille keine Abhängigkeit der Falschrückweisungsrate zu verzeichnen ist. Dies bezieht sich jedoch nur auf den Betrachtungszeitraum, in dem für die jeweilige Referenzbasis optimale Umgebungsbedingungen gewährleistet waren (für ICAO Zyklus 2 und für Template Zyklus 1). Das heißt, dass das Tragen einer Brille keinen signifikanten Einfluss auf die Erkennungsleistung hat, wenn die Umgebungsbedingungen bei den Verifikationen nahezu identisch mit den Bedingungen bei der Erstellung der Referenz (Bildaufnahme bzw. Live-Enrolment) sind. Weichen die Bedingungen jedoch ab, hat dies für Brillenträger wesentliche stärkere Auswirkungen hinsichtlich der Verschlechterung der Erkennungsleistung.<sup>44</sup>

Für die Iriserkennung gibt es eine Abhängigkeit zwischen dem Tragen einer Brille und der Erkennungsleistung. Aus den Untersuchungen wird ersichtlich, dass Brillenträger eine signifikant höhere Falschrückweisungsrate hervorrufen, als bei einer statistischen Unabhängigkeit zu erwarten wäre. Dies liegt vor allem darin begründet, dass Brillen zum Teil sehr starke Störungen in den Aufnahmen der Iris hervorrufen (Spiegelungen, Reflexionen, Überdeckung durch Brillenrand).

#### 7.2.5.8 Bart

Für die Gesichtserkennung wird untersucht, ob sich ein Zusammenhang zwischen dem Tragen eines Barts und der Erkennungsleistung aufzeigen lässt. Der Anteil der Barträger an der Testpopulation ist in Abbildung 102 dargestellt. Es ist zu berücksichtigen, dass der überwiegende Teil der Testteilnehmer keinen Bart trägt.

<sup>44</sup> Eine Analyse der Ergebnisse des Betrachtungszeitraums, in dem für die jeweils andere Referenzbasis optimale Umgebungsbedingungen gewährleistet waren, zeigt, dass hier Brillenträger eine signifikant höhere Falschrückweisungsrate hervorrufen, als bei einer statistischen Unabhängigkeit zu erwarten wäre.



**Abbildung 102: Verteilung der Testpopulation hinsichtlich Merkmal Bart**

Barträger erreichen beim Gesichtserkennungssystem von Cognitec signifikant bessere Erkennungsleistungen, als bei Unabhängigkeit zu erwarten wäre.

#### 7.2.5.9 Zusammenfassung

Eine Übersicht, die darstellt, wo Zusammenhänge zwischen der Erkennungsleistung (FRR) und demographischen bzw. merkmalsbeeinflussenden Eigenschaften bestehen, ist in Abbildung 103 enthalten.

Merkmal	Cognitec	BDr/NEC	Dermalog	SD Industries
<b>Geschlecht</b>	Abhängigkeit mit deutlichem Trend	Abhängigkeit mit deutlichem Trend	Keine Abhängigkeit	Keine Abhängigkeit
<b>Alter</b>	Nicht ermittelbar oder uneinheitliches Verhalten	Abhängigkeit mit deutlichem Trend	Abhängigkeit mit deutlichem Trend	Abhängigkeit mit deutlichem Trend
<b>Ethn. Herkunft</b>	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten
<b>Ausbildung</b>	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten	Nicht ermittelbar oder uneinheitliches Verhalten
<b>Beruf. Tätigkeit</b>	Nicht ermittelbar oder uneinheitliches Verhalten	Abhängigkeit mit deutlichem Trend	Nicht ermittelbar oder uneinheitliches Verhalten	Abhängigkeit mit deutlichem Trend
<b>Kontaktlinsen</b>	Kein Betrachtungsobjekt	Kein Betrachtungsobjekt	Kein Betrachtungsobjekt	Nicht ermittelbar oder uneinheitliches Verhalten
<b>Brille</b>	Keine Abhängigkeit	Kein Betrachtungsobjekt	Kein Betrachtungsobjekt	Abhängigkeit mit deutlichem Trend
<b>Bart</b>	Abhängigkeit mit deutlichem Trend	Kein Betrachtungsobjekt	Kein Betrachtungsobjekt	Kein Betrachtungsobjekt

■ Keine Abhängigkeit      ■ Nicht ermittelbar oder uneinheitliches Verhalten  
■ Abhängigkeit mit deutlichem Trend       Kein Betrachtungsobjekt

**Abbildung 103: Korrelationen zwischen demographischen Merkmalen und Erkennungsleistung**

## 7.2.6 Zeitliche Performance

Bei den im Folgenden angegebenen Performance-Messungen ist stets die in BioP II verwendete Architektur zu berücksichtigen. Aufgrund der Tatsache, dass verschiedene Datensätze (zum Teil inklusive aufgenommener Bilder) in einer Datenbank protokolliert wurden, ergeben sich Beeinflussungen der ermittelten Zeiten. Diese sind jedoch nur dahingehend relevant, dass die ermittelten durchschnittlichen Zeiten um einen annähernd festen Betrag größer sind als bei einer Einsatzumgebung ohne entsprechende Protokollierung. Die direkte Vergleichbarkeit der Zeiten untereinander ist hiervon unberührt.

Nachfolgend werden folgende Ergebnisse zur zeitlichen Performance dargestellt:

- **Verarbeitungszeit:** Als Verarbeitungszeit wird hier die Zeitspanne zwischen dem Erkennen des Tokens durch das biometrische System und dem Vorliegen des Verifikationsergebnisses verstanden. Sie umfasst damit die Aufnahme des biometrischen Merkmals, das Lesen der Referenzen aus der Datenbank, die eigentliche biometrische Verifikation sowie die Protokollierung der Ergebnisse. In der Aufstellung der Verarbeitungszeiten in Tabelle 32 wird in minimale, maximale und mittlere Verarbeitungszeit unterschieden. Da die Verifikation für die sekundäre Referenz im Hintergrund lief, werden nur die Verarbeitungszeiten der Masterreferenz dargestellt.
- **Mittlerer Durchsatz:** Die mittlere Durchsatzrate beschreibt die Anzahl von Benutzern, die pro Stunde durch ein biometrisches System bedient werden können. Als Zeitdauer für die Bedienung eines Benutzers wird dabei die mittlere Verarbeitungszahl zugrunde gelegt.

RefID	Test-zyklus	Verarbeitungszeit [s]			Mittlerer Durchsatz
		minimale	maximale	mittlere	
1	2	4,9	30,4	10,0	362
2	1	3,1	25,9	6,2	580
4	1+2	3,7	63,9	5,9	613
4	1	3,7	18,2	6,2	584
4	2	3,7	63,9	5,8	620
6	1+2	1,6	119,6	4,4	822
6	1	1,6	92,0	4,6	780
6	2	1,6	119,6	4,2	863
8	1+2	2,0	26,0	4,2	850
8	1	2,0	16,0	4,4	810
8	2	2,0	26,0	4,1	888

**Tabelle 32: Verarbeitungszeiten in [s] und mittlerer Durchsatz**

Neben der dargestellten technischen Analyse auf Basis der protokollierten Ergebnisdatensätze wurden innerhalb von Stichproben jeweils zu Beginn eines Feldtestzyklus mittels Stoppuhr die vollständigen Bedienzeiten von Testteilnehmern gemessen. Die Bedienzeiten umfassen neben den oben genannten Verarbeitungszeiten auch den Zeitbedarf für Positionierung vor dem System, Auflegen und Auslesen der Karte sowie visuelles Erfassen des Verifikationsergebnisses durch den Benutzer. Diese Untersuchungen sind im Rahmen der Benutzbarkeitsstudie in Abschnitt 8.1.2.2 dargestellt.

### 7.2.7 Template-Größe

Tabelle 33 gibt einen Überblick über die Größe der beim Live-Enrolment an den biometrischen Systemen erzeugten Templates.

Angaben in [kB]	Cognitec	BDr/NEC	Dermalog	SD Industries
Minimum	4,1	0,1	0,2	0,5
Maximum	10,2	0,2	0,7	0,5
<b>Durchschnitt</b>	<b>8,9</b>	<b>0,1</b>	<b>0,4</b>	<b>0,5</b>

**Tabelle 33: Template-Größen**

Die Iris-Templates von SD Industries haben eine konstante Größe von 512 Byte.

Bei Cognitec ist die Größe des Templates abhängig von der Anzahl der in die Template-Generierung eingehenden Bilder (siehe Tabelle 34).

Anzahl Bilder	Template-Größe [kB]
1	2,0
2	4,1
3	6,1
4	8,1
5 oder mehr	10,2

**Tabelle 34: Template-Größe Cognitec in Abhängigkeit der Anzahl Bilder**

Die Größe der ICAO-Bilder ist in Abschnitt 4.1 im Rahmen der Beschreibung der Erstellung der ICAO-Bilder dargestellt.

### 7.2.8 Allgemeine Ergebnisse zu Systemen und Herstellern

Gemäß Pflichtenheften waren die beteiligten Hersteller aufgefordert, für einen unbetreuten Feldtest geeignete Systeme termingerecht bereitzustellen.

Die beteiligten Systeme und deren Hersteller wurden bezüglich der folgenden Kriterien betrachtet:

- **Inbetriebnahme:** Im Rahmen der Inbetriebnahme sind besonders die termingerechte Bereitstellung der geforderten Funktionen sowie deren korrekte Umsetzung relevant. Die Inbetriebnahme der Systeme erfolgte gemäß Projektplan im Zeitraum 23.02.2004 bis 27.02.2004.
- **Systemfehler:** Dieses Kriterium erfasst, welche Fehler während des Feldtests aufgetreten sind.
- **Ausfallverhalten:** Bei unüberwachtem Betrieb von Systemen ist deren Zuverlässigkeit von hoher Bedeutung. Entsprechend wurden über den Feldtest die Systemausfälle protokolliert.
- **Administrationsaufwand:** Der Einsatz von IT-Systemen erfordert stets einen gewissen Administrationsaufwand. Dieser sollte möglichst gering sein. Des Weiteren sollten geeignete Werkzeuge zur Unterstützung des Administrators bereitstehen, wie beispielsweise Reporting-Mechanismen. Nicht zuletzt ist eine intuitive Bedienung relevant.

- **Service und Support durch die Hersteller:** Unter diesem Kriterium wird zusammengefasst, ob der Hersteller für eine reibungslose Testdurchführung ausreichenden Service und Support geleistet hat. Insbesondere interessiert, ob aufgetretene Fehler gelöst werden konnten.

Auf eine detaillierte Darstellung der Ergebnisse wird in diesem Bericht verzichtet.

## 8 Weiterführende Untersuchungen

Im Rahmen von BioP II wurden neben dem Feldtest eine Reihe weiterführender Untersuchungen durchgeführt. Hierzu zählen:

- **Benutzbarkeit und Akzeptanz:** In einer feldtestbegleitenden Studie wurden die Benutzbarkeit der in BioP II eingesetzten biometrischen Systeme und die Akzeptanz der Benutzer untersucht. Wesentliche Kernpunkte sind in Abschnitt 8.1 dargestellt.
- **Überwindungssicherheit sowie System- und physische Sicherheit:** Auf die detaillierte Darstellung der Ergebnisse wird in diesem Bericht verzichtet.
- **Robustheit und Langlebigkeit der Sensoren:** Die Aufnahmeeinheiten wurden hinsichtlich Abnutzungserscheinungen untersucht (Abschnitt 8.2).
- **Speicherung biometrischer Daten auf RF-Token:** Im Hinblick auf die Verwendung biometrischer Verfahren im Zusammenhang mit Personaldokumenten wird in einem separaten Projekt des BSI die Umsetzung auf Basis von RFID-Technologie untersucht. Für BioP II relevante Kernpunkte dieses Projekts sind in Abschnitt 8.3 dargestellt.

### 8.1 Benutzbarkeit und Akzeptanz

Im Rahmen des Projekts BioP II wurde eine Benutzbarkeits- und Akzeptanzstudie durch das University College London (UCL) durchgeführt. Bestandteile dieser Studie sind eine vergleichende Beurteilung zur Akzeptanz und Benutzbarkeit der in BioP II getesteten Systeme sowie Empfehlungen zur benutzerfreundlichen, akzeptablen und sozialverträglichen Gestaltung biometrischer Anwendungen.

#### 8.1.1 Datenerhebung

Die Teilnahme an der Benutzbarkeits- und Akzeptanzuntersuchung war den Teilnehmern freigestellt. Die Teilnehmer blieben für das Forschungsteam anonym, das heißt es wurde keine Beziehung zu der im Test verwendeten UserID hergestellt. Die Datenerhebung wurde mit den Projektleitern, Datenschutzbeauftragten und Betriebsräten von Fraport and LH abgestimmt. Im Folgenden wird die Datenerhebung im Detail dargestellt:

1. An zwei Tagen der Enrolment-Phase wurden 112 befragungswillige Teilnehmer beobachtet und kurz zu ihren Erfahrungen interviewt.
2. Am Ende der TeachIn-Phase in Testzyklus 1 fanden an drei Tagen 626 Bedienungsbeobachtungen und Kurzbefragungen statt, weitere 353 zu Beginn von Testzyklus 2. Insgesamt wurden 979 Beobachtungen und Kurzbefragungen durchgeführt.
3. Alle Teilnehmer erhielten zwei Fragebögen (jeweils am Ende von Testzyklus 1 und Testzyklus 2), in denen sie die Schnelligkeit und Sicherheit der Systeme sowie die Einfachheit und Bequemlichkeit des Benutzungsvorgangs bewerteten und die benutzten Systeme nach ihrer persönlichen Präferenz ordnen konnten. Außerdem wurde die Zustimmung oder Ablehnung dieser Verfahren für drei Anwendungsszenarien (Geldautomat, PC, Reisepass) erfragt. Am Ende von Zyklus 1 wurden 907 Fragebögen ausgefüllt, am Ende von Testzyklus 2 gab es einen Rücklauf von 568 Fragebögen (insgesamt 1.475 Fragebögen).
4. Alle Teilnehmer hatten weiterhin die Möglichkeit, an Fokusgruppen teilzunehmen, in denen sie ihre Erfahrungen mit den Systemen austauschen und die Vor- und Nachteile von biometrischen Verfahren mit anderen Teilnehmern diskutieren konnten. In Testzyklus 1 wurden 34 Fokusgruppen mit jeweils vier bis acht Teilnehmern über vier Tage durchgeführt

und 23 weitere Fokusgruppen in Testzyklus 2 (insgesamt 57 Fokusgruppen mit über 400 Teilnehmern).

5. Angestellte, die nicht am Feldtest teilgenommen hatten, konnten sich durch einen anonymen Fragebogen zu biometrischen Verfahren und deren Einsatz äußern. Von über 10.000 an die Belegschaft versendeten Fragebögen wurden 647 ausgefüllt.

Die Bewertung der Benutzbarkeit der Systeme (siehe Abschnitt 8.1.2) beruht auf den folgenden Erhebungen:

- Bewertungen der Systeme durch die Teilnehmer und explizit vermerkte Benutzungsprobleme. Datenbasis hierfür sind die Fragebögen aus Testzyklus 1 und Testzyklus 2 (siehe Punkt 3 oben).
- Durch das Forschungsteam erhobene Daten zur Bedienzeit und beobachtete Benutzungsprobleme. Datenbasis hierfür sind die Beobachtungen und Messungen in Testzyklus 1 und Testzyklus 2 (siehe Punkt 2 oben).

Die anderen in dieser Benutzbarkeits- und Akzeptanzstudie erhobenen Daten geben zusätzliche Informationen über die Erfahrungen, Wahrnehmungen und Einstellungen der Teilnehmer und wurden in der Analyse zur Akzeptanz der getesteten Systeme und weiterer Anwendungsszenarien (siehe Abschnitt 8.1.3) herangezogen.

## 8.1.2 Benutzbarkeit

### 8.1.2.1 Subjektive Einschätzungen

Von den Testteilnehmern wurde eine **Gesamtbeurteilung** der Systeme erfragt. Diese ist eine subjektive Einschätzung der Zwecktauglichkeit und des Benutzungskomforts der vier Systeme im direkten Vergleich. Die Erhebung erfolgte mittels eines Fragebogens am Ende von Testzyklus 2. Insgesamt wurden 568 Fragebögen zurückerhalten. In dem Fragebogen sollten die Testteilnehmer ein **Ranking** der Systeme angeben, keine Benotung. Eine „1“ bedeutet, dass der Teilnehmer das System als das Beste einschätzt, während eine „4“ bedeutet, dass dies nach Ansicht des Teilnehmers im direkten Vergleich als das Schlechteste anzusehen ist. Dies führte zu dem in Tabelle 35 dargestellten Ergebnis.

		Cognitec	BDr/NEC	Dermalog	SDI
N	Gültig	425	422	416	435
	Fehlend	49	52	58	39
<b>Mittelwert</b>		<b>2,84</b>	<b>2,27</b>	<b>2,03</b>	<b>1,72</b>
Medianwert		3	2	2	1
Standardabweichung		1,10	0,98	1,00	0,91
Minimum		1	1	1	1
Maximum		4	4	4	4

**Tabelle 35: Durch Teilnehmer vergebene Rangfolge der Systeme**

Die Gesamteinschätzung der Systeme im direkten Vergleich ergibt, dass SD Industries im Durchschnitt als bestes System beurteilt wird, gefolgt von Dermalog und Bundesdruckerei/NEC. Cognitec wird als am schlechtesten eingeschätzt.

In **Einzeleinschätzungen der Benutzungskriterien** Sicherheit, Schnelligkeit, Einfachheit und Bequemlichkeit konnten die Teilnehmer Noten von 1-5 für jedes System vergeben. Die Erhebung erfolgte mittels Fragebogen jeweils am Ende der beiden Testzyklen. Die Teilnehmer wurden dabei gebeten, die zwei in dem betreffenden Zyklus hauptsächlich benutzten Biometriesysteme zu bewerten. Da Teilnehmer die in den Teststationen gepaarten Systeme – das heißt entweder Dermalog und SD Industries oder Bundesdruckerei/NEC und Cognitec – bewerteten, ist nur ein direkter Vergleich zwischen je zwei Systemen möglich. Direkte Vergleiche zwischen allen Systemen wurden von Teilnehmern nicht explizit durchgeführt. Daher sind keine eindeutigen Aussagen möglich, die nicht-gepaarte Systeme (zum Beispiel Dermalog und Bundesdruckerei/NEC) betreffen. Die Noten entsprechen den in Tabelle 36 dargestellten Bezeichnungen.

Note	Sicherheit	Schnelligkeit	Einfachheit	Bequemlichkeit
1	Absolut sicher	Immer schnell genug	Keine Probleme	Erfordert keinen Aufwand
2	Sehr sicher	Meistens schnell genug	Nur selten Probleme	Erfordert geringen Aufwand
3	Sicher genug	Manchmal nicht schnell genug	Manchmal Probleme	Erfordert akzeptablen Aufwand
4	Sollte sicherer werden	Sollte schneller werden	Sollte einfacher werden	Sollte bequemer werden
5	Müsste dringend sicherer werden	Müsste dringend schneller werden	Müsste dringend einfacher werden	Müsste dringend bequemer werden

**Tabelle 36: Klassifikation Benutzungskriterien**

Die durchschnittliche Bewertung der Systeme nach diesen Kriterien ist in Tabelle 37 zusammengefasst.

System	Sicherheit	Schnelligkeit	Einfachheit	Bequemlichkeit
Cognitec	2,29	2,87	1,45	1,61
BDr/NEC	2,25	2,38	1,22	1,39
Dermalog	2,35	2,20	1,40	1,40
SD Industries	2,06	1,79	1,34	1,47

**Tabelle 37: Mittlere Bewertung nach Benutzungskriterien<sup>45</sup>**

Für *Sicherheit* liegen die Bewertungen zwischen „Sehr sicher“ und „Sicher genug“. In der Bewertung der *Schnelligkeit* sind die häufigsten Nennungen „Meistens schnell genug“ und „Manchmal nicht schnell genug“. Für *Einfachheit* und *Bequemlichkeit* vergeben die meisten Teilnehmer die besten Bewertungen, die den Bezeichnungen „Keine Probleme“ und „Nur selten Probleme“ bzw. „Erfordert keinen Aufwand“ und „Erfordert geringen Aufwand“ entsprechen.

Da alle Teilnehmer diese Bewertungen jeweils nach Ende eines Testzyklus für die zwei in diesem Zyklus getesteten Systeme durchführten, kann ein statistischer Vergleich der Mittelwerte nur gepaart stattfinden. SD Industries wird bei der Sicherheit, Schnelligkeit und Einfachheit besser als Dermalog

<sup>45</sup> Für die Systembewertung anhand einer Skala von 1 bis 6 (siehe Kapitel 9) werden diese Werte gemäß der folgenden Berechnungsvorschrift umskaliert:  $\frac{\text{Mittelwert}}{5} * 6$ .

eingeschätzt, während Dermalog in der Einschätzung der Bequemlichkeit besser als SDI ist. Die Teilnehmer beurteilen Bundesdruckerei/NEC besser als Cognitec für Schnelligkeit, Einfachheit und Bequemlichkeit. Im Punkt der Sicherheit werden BDr/NEC und Cognitec etwa gleichgestellt.

Ähnlich wie bei den Ergebnissen des Feldtests, die einen Zusammenhang zwischen Nutzungshäufigkeit und Erkennungsleistung feststellen (siehe Abschnitt 7.2.3.2), zeigen sich Unterschiede in der subjektiven Beurteilung der Systeme zwischen Teilnehmern, die nach eigenen Angaben die Systeme regelmäßig oder nicht regelmäßig benutzten. Unregelmäßige Benutzer bewerten das GE-System in der Gesamteinschätzung besser als regelmäßige Benutzer. Für die FE-Systeme und das IE-System vergeben regelmäßige Benutzer insgesamt bessere Noten für die Kriterien Schnelligkeit, Sicherheit, Einfachheit und Bequemlichkeit als unregelmäßige Benutzer.

### 8.1.2.2 Objektive Daten

Die Erhebung der **Bedienzeiten** erfolgte durch Beobachtungen von Betätigungen zufällig ausgewählter Teilnehmer jeweils zu Beginn der beiden Testzyklen (zum Teil während der TeachIn-Phase). Die manuell gestoppten Bedienzeiten weichen erheblich von den im Feldtest gemessenen Verarbeitungszeiten ab (siehe Abschnitt 7.2.6). Die Bedienzeit umfasst nicht nur die vom System protokollierte Verarbeitungszeit, sondern die Zeit für den kompletten Bedienungsvorgang vom Auflegen des RF-Token bis zum Verlassen des Gerätes. Bei der über beide Beobachtungsphasen gemittelten Bedienzeit ist das GE-System mit 11,1 Sekunden am schnellsten, die FE-Systeme liegen mit 12,1 Sekunden (Dermalog) und 12,6 Sekunden (BDr/NEC) im Mittelbereich, während das IE-System mit durchschnittlich 13,3 Sekunden die längste Bedienzeit hat. Die Unterschiede zwischen der subjektiv erfahrenen Bedienzeit der Teilnehmer (siehe Abschnitt 8.1.2.1) und der objektiv beobachteten Bedienzeit kann damit erklärt werden, dass Wartezeiten ohne Rückmeldung oft als länger empfunden werden. Die akustische Benutzerführung und Rückmeldung des IE-Systems lässt den Benutzungsvorgang kürzer erscheinen, während beim GE-System längere Wartezeiten ohne Benutzeraktivität anfallen. Die objektiv lange Bedienzeit für das IE-System ist durch Benutzbarkeitsprobleme (siehe unten) zu erklären, die vielen unregelmäßigen Benutzern (aber auch einer Teilgruppe von regelmäßigen Benutzern) massive Schwierigkeiten bereiteten und den Benutzungsvorgang verlängerten.

Bei der Untersuchung der **Benutzbarkeitsprobleme** wurden zwei grundsätzliche Bereiche identifiziert, die die Benutzbarkeit beeinträchtigten:

1. Die Teilnehmer hatten Schwierigkeiten, sich optimal an den Geräten zu positionieren und/oder das relevante biometrische Merkmal richtig zu präsentieren.
2. Mangelndes ergonomisches Design, Gestaltung der Benutzeroberfläche sowie fehlende oder inadäquate Benutzerführung und Rückmeldung führen zu Bedienungsfehlern.

Die Bewertung der Benutzbarkeitsprobleme beruht auf den von den Teilnehmern in den Fragebögen explizit berichteten Bedienungsschwierigkeiten.

- **Schwierigkeiten der Positionierung**

- **Cognitec:** Beim Cognitec-System gab es im Wesentlichen nur Probleme bei der Einhaltung eines geeigneten Abstands zum System. Die am Boden angebrachten Markierungen wurden von vielen Teilnehmern nicht beachtet.
- **Bundesdruckerei/NEC und Dermalog:** Bei den Fingerabdruckererkennungssystemen entstanden Probleme bei der Positionierung durch die Auflage des Fingers auf den Sensoren. Verschiedene Teilnehmer wussten nicht wie der Finger aufzulegen war (Höhe und Druck) und einige Frauen mit langen Fingernägeln beschwerten sich über die unbändige Auflage des Dermalog-Gerätes.

- **SD Industries:** Bei SDI traten Positionierungsschwierigkeiten gehäuft auf, da Teilnehmer gleichzeitig den richtigen Abstand zum Gerät und den Fokuspunkt (Positionierung des Auges in der Mitte des Spiegels der Sensoreinheit) finden mussten. Teilnehmer mussten sich oft ein Auge zuhalten, um zu fokussieren. Die Kopplung mit einer trägen akustischen Rückmeldung bzgl. des Abstands führte zu „Schaukel“-Bewegungen. Außerdem wurde von einer Reihe von Teilnehmern die Sensoreinheit nicht auf die der Körpergröße angemessene Position eingestellt. Viele größere Teilnehmer bückten sich, anstatt die Einstellung zu verändern.
- **Mangelndes ergonomisches Design**
  - **SD Industries:** Die schräge Auflage für das RF-Token führte bei einigen Teilnehmern zu Schwierigkeiten, vor allem wenn sie sich ein Auge zuhalten mussten, um zu fokussieren. Weiterhin ist der Einstellbereich der Sensoreinheit für sehr klein- oder großwüchsige Personen nicht ausreichend.
- **Mangelnde Benutzerführung / Rückmeldung**

Wartezeiten von mehr als 2 Sekunden, in denen die Benutzer keine Rückmeldung erhalten, werden oft als länger empfunden und können zu Irritation und unkooperativem Benutzerverhalten führen.

  - **Cognitec:** Die Rückmeldung ist nur in der Form von kleinen „Ampel“-Leuchtdioden gegeben. Es besteht keine weitergehende Benutzerführung.
  - **Bundesdruckerei/NEC:** Auf einem Monitor gibt es detaillierte Anweisungen. Jedoch waren Teilnehmer unsicher, wann sie den Finger auflegen dürfen, da die Zeitspanne für das Erkennen und Lesen des RF-Token relativ lange dauerte. Außerdem muss der Benutzer zwischen Monitor und RF-Leser bzw. Sensor hin- und herblicken, was den Vorgang verlangsamt.
  - **Dermalog:** Das Dermalog-System hat eine minimale Oberflächengestaltung. Ein Vorteil dieser Benutzeroberfläche ist, dass alle notwendigen Bedienelemente nah zusammen und im Blickfeld des Benutzers sind. Der Nachteil ist, dass Teilnehmer anfangs nicht wissen, was die Leuchtdioden bedeuten. Farbenblinde Teilnehmer hatten bei diesem System die meisten Schwierigkeiten, die Rückmeldung richtig zu interpretieren.
  - **SD Industries:** Obwohl SD Industries die umfangreichste Benutzerführung und Rückmeldung aufweist, empfinden einige Teilnehmer die Stimme auf Dauer als lästig. Auch kommt die Benutzerführung, vor allem zum Abstand des Teilnehmers zum Gerät, zu langsam. Daher kommt es zu „Schaukel“-Bewegungen.

Die in Tabelle 38 angegebene Anzahl und die Prozentsätze beziehen sich auf explizit erwähnte Probleme über die erfassten Teilnehmerberichte (1.475 Fragebögen).

Problemkategorie	Cognitec		BDr/NEC		Dermalog		SDI	
	Anz.	[%]	Anz.	[%]	Anz.	[%]	Anz.	[%]
Positionierung	24	1,63	21	1,42	26	1,76	77	5,22
Ergonomie	0	0,00	0	0,00	0	0,00	28	1,90
Benutzerführung / Rückmeldung	22	1,49	12	0,81	17	1,15	19	1,29

**Tabelle 38: Benutzbarkeitsprobleme**

### 8.1.3 Akzeptanz

Zur Akzeptanz der getesteten Systeme wurden die Teilnehmer gefragt, ob sie dem Einsatz von Biometrie in den drei Anwendungsszenarien PC, Geldautomat und Reisepass zustimmen würden. Die Erhebung erfolgte mittels Fragebogen jeweils am Ende der beiden Testzyklen. Für die Klassifikation der Akzeptanz wurde eine 5-Punkte Skala mit den folgenden Bezeichnungen verwendet:

- 1 Bin dafür
- 2 Eher positiv
- 3 Neutral
- 4 Eher negativ
- 5 Bin dagegen

Die Mittelwerte der Antworten sind in Tabelle 39 dargestellt.

System	PC	Geldautomat	Reisepass
Cognitec	2,29	2,05	1,87
BDr/NEC	1,78	1,83	1,71
Dermalog	1,97	2,06	1,85
SD Industries	2,08	1,77	1,64

**Tabelle 39: Mittlere Bewertung der Akzeptanz nach Anwendungsszenarien**

Die Antworten der Teilnehmer lagen für alle Anwendungsszenarien im positiven Bereich, den Bewertungen „Bin dafür“ oder „eher positiv“ entsprechend. Der niedrigste Mittelwert (für den Einsatz des GE-System am PC) lag bei 2,3. Der Einsatz von Biometrie im Reisepass erhielt die größte Zustimmung.

Trotz mehrheitlicher Befürwortung des Einsatzes biometrischer Systeme in den betreffenden Anwendungsszenarien hatten Teilnehmer in den Fokusgruppen viele Fragen oder Bedenken zur Zuverlässigkeit, Überwindungssicherheit und gesundheitlichen Unbedenklichkeit der getesteten Systeme sowie zum Schutz der biometrischen Daten. Dabei traten insbesondere zwei Bereiche auf:

- Trotz der Befürwortung des Anwendungsszenarios „Reisepass“ durch die Teilnehmer in den Fragebögen wurden in fast allen Fokusgruppen Datenschutzbedenken im Zusammenhang mit grenzüberschreitendem Verkehr geäußert. Die meisten Teilnehmer waren unsicher, ob biometrische Daten beim Grenzübertritt kopiert und eventuell von anderen Staaten verwendet werden können.
- Trotz der positiven Bewertung des IE-Systems durch die Teilnehmer in den Fragebögen ist zu vermerken, dass die meisten der von den Teilnehmern geäußerten Fragen und Bedenken auf dieses System gerichtet sind. Hier stehen gesundheitliche Unbedenklichkeit und nicht-zweckgebundene Anwendung (zum Beispiel Gesundheitsdiagnose) im Vordergrund.

### 8.1.4 Empfehlungen zur Gestaltung der Systeme

Die beobachteten Benutzungsvorgänge, Benutzungsprobleme und Aussagen der Teilnehmer in Fragebögen und Fokusgruppen bilden die Grundlage der Empfehlungen zur Verbesserung der Benutzbarkeit der Systeme.

Für alle Systeme ist festzuhalten, dass in einer realen Einsatzumgebung die Gestaltung des Feedbacks durch die Anwendung unterstützt wird, zum Beispiel das Öffnen einer Schranke bei einer Zutrittskontrolle.

Beim **GE-System von Cognitec** bedarf die Benutzeroberfläche dringend einer Überarbeitung. Die jetzige Gestaltung genügt den ästhetischen Ansprüchen der Teilnehmer nicht und unterminiert das Vertrauen in die technische Zuverlässigkeit des Systems. Benutzerführung und Rückmeldung sind nahezu nicht vorhanden. Die folgenden Änderungen könnten zu einer Verbesserung führen:

1. Eine Zielbildanzeige, in der der Benutzer sehen kann, wie er von der Kamera erfasst wird.
2. Die Ampel-Rückmeldung ist zu klein, uninteressant und für Farbenblinde ungeeignet. Zu empfehlen ist eine gut sichtbare und visuell attraktivere Anzeige, die die Aufmerksamkeit des Benutzers auf sich zieht.
3. Es sollte eine gezielte Rückmeldung - zum Beispiel „bitte still stehen“, „bitte hierhin schauen“ – durch Anzeige, Sprachrückmeldung oder eine Kombination von beidem erfolgen. Bei Verwendung eines akustischen Feedbacks gelten die Empfehlungen für die Gestaltung der Sprachrückmeldung wie beim IE-System (siehe unten).

Beide **FE-Systeme** haben mit Bedenken hinsichtlich Hygiene zu kämpfen. Kontaktlose Sensoren könnten die Akzeptanz dieser Systeme erhöhen, außerdem wäre die Genauigkeit bei der Positionierung evtl. weniger kritisch. Auch Probleme mit mangelndem oder zu viel Druck sowie mit zu viel oder zu wenig Hautfeuchtigkeit könnten damit ggf. vermieden werden.

Das System von **Bundesdruckerei/NEC** gibt detaillierte Anweisungen und Rückmeldung. Diese benutzerfreundliche Gestaltung des Systems kam jedoch durch mangelnde Schnelligkeit nicht zum Tragen. Das System kann wie folgt verbessert werden:

1. Für geübte Benutzer sind detaillierte Anweisungen nicht notwendig.
2. Für unregelmäßige Benutzer könnten die Anweisungen noch wie folgt verbessert werden:
  - a. Anzeige einer Hand, mit gewünschtem Finger markiert.<sup>46</sup>
  - b. Anzeige/Anweisung, den Finger gerade und flach aufzulegen.
3. Ein an der Kante des Gerätes angebrachter Sensor könnte die Auflage des relevanten Teils des Fingers erleichtern.

Das System von **Dermalog** hat eine kompakte Benutzeroberfläche, die für geübte Nutzer gut funktioniert. Für unregelmäßige Benutzer ist die Anweisung und Rückmeldung nicht ausreichend. Die folgenden Änderungen könnten zu einer Verbesserung führen:

1. Die Ampel-Rückmeldung ist zu klein und für Farbenblinde ungeeignet. Die Rückmeldung mittels einer Anzeige mit Symbolen (zum Beispiel +/- oder ☺/☹) könnte hier Abhilfe schaffen.
2. Ein bündiger Sensor, der die Auflage von Fingern mit langen und stark gebogenen, unflexiblen Nägeln nicht unnötig erschwert, sollte eingesetzt werden.
3. Es sollte genauere Anweisungen für unregelmäßige Benutzer geben.

Beim **IE-System von SD Industries** ist die jetzige Höhenverstellung nicht für Rollstuhlfahrer und kleinere oder größere Benutzer geeignet. Das System – oder eine Installation mehrerer Systeme – muss so konfiguriert werden, dass alle Benutzer das System erreichen können. Weiterhin wird die schon

---

<sup>46</sup> Dies war im Feldtest von BioP II ausgeschlossen.

jetzt mögliche Höheneinstellung oft nicht wahrgenommen oder von den Teilnehmern gemieden (zum Beispiel aus hygienischen Gründen). Die Probleme bei der Fokussierung werden bei Systemen, die beide Augen erfassen, vermieden. Jedoch haben Benutzer mit unterschiedlicher Fokussierung auf beiden Augen oft Schwierigkeiten mit solchen Systemen. Folgende Vorschläge könnten zu einer Verbesserung führen:

1. Ein Gerät für die Aufnahme nur eines Auges könnte durch bessere visuelle Rückmeldung der erwarteten und tatsächlichen Augenposition verbessert werden. Noch einfacher wäre ein System mit automatischer Höhenverstellung und Autofokus. Dies hat jedoch Implikationen für die Größe der Systeme und die Herstellungskosten.
2. Die Fläche des Kartenlesers sollte waagrecht sein, so dass die Karte abgelegt werden kann.
3. Die Rückmeldung über Positionierungsprobleme sollte schneller erfolgen. Die akustische Rückmeldung sollte auf Wunsch abschaltbar sein. Dies ist besonders für regelmäßige Benutzer wichtig.

## 8.2 Robustheit und Langlebigkeit der Sensoren

Die eingesetzten Sensoren wurden auf ihre Langzeitstabilität untersucht. Durch den eingeschränkten Testzeitraum von fünf Monaten (inkl. Enrolment) können nur bedingt Rückschlüsse gezogen werden. Für kontaktbehaftete Sensoren (im Wesentlichen die Fingerabdrucksensoren, aber auch die Kameras zur Iriserkennung aufgrund mechanischer Höheneinstellung) sind in dieser Zeit jedoch viele verschleißbehaftete Betätigungen erfolgt. Eine entsprechende Übersicht kann Tabelle 40 entnommen werden. Die tatsächliche Anzahl liegt etwas höher, da die Betätigungen der Inbetriebnahme- und Kalibrierungsphase (vor dem 12.03.2004) nicht berücksichtigt sind.

System	Betätig.	System	Betätig.	System	Betätig.	System	Betätig.
COG02	12.584	BDR02	17.935	DER02	16.517	SDI02	13.472
COG03	14.837	BDR03	14.791	DER03	14.435	SDI03	19.541
COG04	17.766	BDR04	12.163	DER04	12.302	SDI04	17.931
COG05	12.941	BDR05	16.514	DER05	18.087	SDI05	13.409
COG06	10.163	BDR06	14.610	DER06	13.009	SDI06	8.977
COG07	10.485	BDR07	11.825	DER07	9.885	SDI07	10.088
COG08	13.079	BDR08	12.716	DER08	9.945	SDI08	12.615
COG09	15.482	BDR09	13.152	DER09	10.301	SDI09	14.542

**Tabelle 40: Anzahl Betätigungen an den biometrischen Sensoren**

Alle Sensoren wurden hinsichtlich der folgenden Kriterien untersucht:

- **Qualität der Aufnahme:** Visueller Vergleich von Merkmalsaufnahmen bei Inbetriebnahme und nach Ende des Feldtests zwecks Analyse, ob es zu einer Verschlechterung der Bildqualität gekommen ist
- **Verschleiß:** Untersuchung von Sensorfläche und Sensorgehäuse hinsichtlich Gebrauchs- und Verschleißspuren

## 8.2.1 GE-System von Cognitec

Das Cognitec-System verwendet vier USB-Kameras des Typs „Philips ToUcam Pro, PCVC 740K“ zur Bildaufnahme.

### 8.2.1.1 Qualität der Aufnahme

Seitens des Herstellers wurde der Abstand zwischen der Person und dem Gerät auf etwa 25cm festgelegt. Dies führt zu einem Gesamtabstand zwischen Kamera und Gesicht von etwa 50cm.

System	Inbetriebnahme	Ende Feldtest
COG02	OK	OK
COG03	Kamera 2 unscharf	Kamera 2 unscharf
COG04	OK	OK
COG05	Kamera 1 unscharf	Kamera 1 unscharf
COG06	OK	OK
COG07	OK	OK
COG08	Kamera 1 unscharf Kamera 3 unscharf Kamera 4 unscharf	Kamera 1 unscharf Kamera 3 unscharf Kamera 4 unscharf
COG09	OK	OK

**Tabelle 41: Aufnahmequalität Cognitec**

Eine Verschlechterung der Bildqualität konnte nicht festgestellt werden. Die Bildqualität war über den gesamten Testzeitraum konstant.

### 8.2.1.2 Verschleiß

An den Sensoren konnten keinerlei Gebrauchs- und Verschleißspuren gefunden werden.

Die Sensoren sind nicht gegen Verschmutzung geschützt. Im Gegensatz zu den anderen Biometricsystemen können die Sensoren (Kameras) nicht gereinigt werden. Ablagerungen durch Luftverschmutzung, wie zum Beispiel Zigarettenqualm, können sich auf der Sensoroptik niederschlagen.

## 8.2.2 FE-System von Bundesdruckerei/NEC

Das BDr/NEC-System verwendet den Fingerabdruck-Scanner „ACCO 1394“ von Smiths Heimann Biometrics zur Bildaufnahme.

### 8.2.2.1 Qualität der Aufnahme

Zur Bewertung der Qualität der Aufnahme wird in zwei Kategorien unterschieden.

- **Blankobilder**

Die Blankobilder sollen zeigen, ob Defekte oder Verschmutzungen innerhalb des Sensors existieren.

System	Inbetriebnahme	Ende Feldtest
BDR02	OK	OK
BDR03	OK	OK
BDR04	OK	OK
BDR05	OK	OK
BDR06	OK	OK
BDR07	OK	OK
BDR08	OK	OK
BDR09	OK	OK

**Tabelle 42: Aufnahmequalität Blankobilder bei Bundesdruckerei/NEC**

Es konnten keinerlei Defekte oder Verunreinigungen festgestellt werden.

- **Fingerabdruckbilder**

In diesem Bereich werden die Bilder hinsichtlich Qualität, Schärfe und Kontrast untersucht.

System	Inbetriebnahme	Ende Feldtest
BDR02	OK	Grauschleier
BDR03	OK	OK <sup>47</sup>
BDR04	OK	OK
BDR05	OK	OK
BDR06	OK	OK
BDR07	OK	OK
BDR08	OK	OK
BDR09	OK	Grauschleier

**Tabelle 43: Aufnahmequalität Fingerabdruckbilder bei Bundesdruckerei/NEC**

Zum Ende des Tests lieferten zwei Sensoren Bilder mit einem Grauschleier. Bei allen anderen Sensoren konnten keinerlei qualitative Unterschiede in den Bildern gefunden werden.

<sup>47</sup> Auch hier trat der Grauschleier auf, jedoch erst nach Ende des Feldtests.



**Abbildung 104: Fingerabdruckbild links ohne und rechts mit Grauschleier bei BDR09<sup>48</sup>**

Eine detaillierte Analyse ergab, dass bei den Systemen BDR02 und BDR09 der Fehler ab dem 01.07.2004 auftrat.

Zum Zeitpunkt des Fehlers war das Gerät BDR09 bei Fraport im Vorstandsgebäude aufgestellt. Vor der direkten Sonneneinstrahlung wurde das Gerät durch einen teillichtdurchlässigen Vorhang geschützt. Die Geräte BDR02 und BDR03 waren bei Lufthansa am Aufstellort Tor 21 aufgestellt. Durch bauseitig aufgestellte Trennwände waren die Geräte vor der Sonneneinstrahlung geschützt.

Auf die Erkennungsleistung hatte dieses Verhalten keinen messbaren Einfluss. Die Falschrückweisungsrate der betroffenen Systeme lagen auch nach dem 01.07.2004 in einem ähnlichen Bereich wie die der anderen Systeme.

#### 8.2.2.2 Verschleiß

An den Sensorflächen und den Sensorgehäusen konnten keinerlei Gebrauchs- und Verschleißspuren gefunden werden.

Die Temperatur des Sensorgehäuses steigt bei höheren Umgebungstemperaturen überproportional an.

#### 8.2.3 FE-System von Dermalog

Das Dermalog-System verwendet den Fingerabdruck-Scanner „CrossMatch Verifier 300“ zur Bildaufnahme.

Die Art der Beleuchtung innerhalb des Sensors, lässt die Glasfläche schnell schmutzig erscheinen. Aufgrund der regelmäßigen Reinigung der Sensoroberfläche hat dies jedoch offenbar keinen wesentlichen Einfluss auf die Erkennungsleistung.

##### 8.2.3.1 Qualität der Aufnahme

Zur Bewertung der Qualität der Aufnahme wird in zwei Kategorien unterschieden.

- **Blankobilder**

Die Blankobilder sollen zeigen, ob Defekte oder Verschmutzungen innerhalb des Sensors existieren.

---

<sup>48</sup> Zur besseren Veranschaulichung wurde in diesem Bild der Kontrast gegenüber dem Originalbild verstärkt.

System	Inbetriebnahme	Ende Feldtest
DER02	0 Pixelfehler	0 Pixelfehler
DER03	2 Pixelfehler	3 Pixelfehler
DER04	3 Pixelfehler	3 Pixelfehler
DER05	3 Pixelfehler	3 Pixelfehler
DER06	4 Pixelfehler	4 Pixelfehler
DER07	1 Pixelfehler	1 Pixelfehler
DER08	2 Pixelfehler	2 Pixelfehler
DER09	0 Pixelfehler	0 Pixelfehler

**Tabelle 44: Aufnahmequalität Blankobilder bei Dermalog**

Die Ausleuchtung der Sensorfläche ist in den Aufnahmen schwach zu erkennen.

Das System DER03 brachte ab dem 03.05.2004 einen weiteren Pixelfehler.

- **Fingerabdruckbilder**

In diesem Bereich werden die Bilder hinsichtlich Qualität, Schärfe und Kontrast untersucht.

System	Inbetriebnahme	Ende Feldtest
DER02	OK	OK
DER03	OK	OK
DER04	OK	OK
DER05	OK	OK
DER06	OK	OK
DER07	OK	OK
DER08	OK	OK
DER09	OK	OK

**Tabelle 45: Aufnahmequalität Fingerabdruckbilder bei Dermalog**

Die Bildqualität war über den gesamten Testzeitraum konstant.

### 8.2.3.2 Verschleiß

An den Sensorflächen und den Sensorgehäusen konnten keinerlei Gebrauchs- und Verschleißspuren gefunden werden.

## 8.2.4 IE-System von SD Industries

Das SDI-System verwendet den Sensor „IrisAccess 3000“ der Firma LG Electronics zur Bildaufnahme.

### 8.2.4.1 Qualität der Aufnahme

In diesem Bereich werden die Bilder hinsichtlich Qualität, Schärfe und Kontrast untersucht.

System	Inbetriebnahme	Ende Feldtest
SDI02	OK	OK
SDI03	OK	OK
SDI04	OK	OK
SDI05	OK	OK
SDI06	OK	OK
SDI07	OK	OK
SDI08	OK	OK
SDI09	OK	OK

**Tabelle 46: Aufnahmequalität SD Industries**

Die Qualität der Aufnahmen war über den gesamten Testzeitraum stabil.

#### 8.2.4.2 Verschleiß

An Sensoroptik, Sensorgehäuse und Mechanik konnten keinerlei Gebrauchs- und Verschleißspuren gefunden werden.

#### 8.2.5 Zusammenfassung

System	Qualitätseinbußen	Physischer Verschleiß
Cognitec	Keine Einbußen	Kein Verschleiß
BDr/NEC	Grauschleier an zwei Sensoren	Kein Verschleiß
Dermalog	Vereinzelte Pixelfehler	Kein Verschleiß
SD Industries	Keine Einbußen	Kein Verschleiß

**Tabelle 47: Zusammenfassung Sensorrobustheit**

### 8.3 Speicherung biometrischer Daten auf RF-Token

Aufgrund der Fokussierung von BioP auf die Verwendung biometrischer Systeme im Zusammenhang mit Personaldokumenten berücksichtigt das Projektkonzept wesentliche damit einhergehende Anforderungen. So wird als Verifikationsmittel ein kontaktloses RF-Token (Philips mifare 1k) als Simulation eines Chips in einem Personaldokument verwendet. Aufgrund der zum Projektstart vorliegenden Verfügbarkeit und der Kostensituation wurden keine leistungsfähigeren Pendants eines elektronischen Ausweises verwendet. Stattdessen simuliert die Kombination aus RF-Token und referenzierter Datenbank mit standard-konformer Speicherung und Kodierung der biometrischen Merkmale ein elektronisches Personaldokument gemäß ICAO. Mittels der durch secunet bereitgestellten Schnittstelle zwischen biometrischem System und Verifikationsmittel (RF-API) können jedoch auch leistungsfähige kontaktlose Smartcards angesprochen werden, die direkt die biometrischen Merkmale enthalten. Gemäß den Pflichtenheften für die Hersteller sind auch deren Systeme entsprechend vorbereitet.

Da über die Projektlaufzeit mehrere Anpassungen der relevanten Standards der ICAO erfolgten, wurden die vorgesehenen Untersuchungen in eine separate Projektreihe des BSI ausgelagert, die aktiv

die Standardisierungsmaßnahmen der ICAO begleitet. Diese Projektreihe wird analog zu BioP II in Zusammenarbeit von BSI, BKA und secunet durchgeführt.

Nachfolgend ist ein Auszug von bereits vorliegenden Projektergebnissen dargestellt, der wesentliche Fragestellungen aus BioP II beantwortet.

- **Realisierbarkeit biometrische Daten auf RF-Token:** Auf Basis vorliegender Lösungen kann grundsätzlich die Speicherung biometrischer Daten auf einer kontaktlosen Smartcard erfolgen. Es gibt jedoch gewisse technische Einschränkungen wie beispielsweise der verfügbare Speicherplatz und mögliche Übertragungszeiten. Die untersuchten Smartcards bieten eine effektive Speicherkapazität von maximal 64 kB. Es gibt jedoch Karten, die mehr Speicherplatz bieten, beispielsweise eine Sharp-Lösung mit bis zu 1 MB.
- **Mögliche kontaktlose Smartcards und Betriebssysteme:** Derzeit wird von nahezu allen europäischen Herstellern der Chip Philips SmartMX als Basis genutzt. Infineon konnte bis zum ersten Quartal 2005 keinen serientauglichen Chip mit gleicher Kapazität aufweisen. Erste Prototypen sind allerdings bereits im Umlauf. Folgende Smartcard-Betriebssysteme wurden auf den SmartMX bereits aufgesetzt und haben ihre Tauglichkeit im Biometrie-Umfeld bewiesen:
  - EvalOS: Es handelt sich hierbei um eine prototypische Implementierung von Philips, die nicht weitergepflegt wird. EvalOS bietet keine kryptografische Funktionalität.
  - StarCOS 3.0: Dieses Betriebssystem von Giesecke & Devrient ist serienreif und verfügt über eine ausgeprägte kryptografische Funktionalität. StarCOS 3.0 erfüllt die Anforderungen der LDS-Spezifikation sowie der aktuellen PKI-Spezifikationen.
  - JCOP 3.0 mit Applet der Bundesdruckerei: Diese Lösung ist serienreif und stellt eine spezialisierte kryptografische Funktionalität bereit. Sie ist speziell abgestimmt auf die LDS-Spezifikation sowie die aktuellen PKI-Spezifikationen.
- **Übertragungsgeschwindigkeiten:** Die Performance bei der kontaktlosen Datenübertragung ist im Wesentlichen vom verwendeten Leser, dem Betriebssystem der Karte sowie der Struktur der Daten abhängig. Java-Applets auf Karten sind vor allem bei der gesicherten Übertragung von Daten wesentlich langsamer als native Betriebssysteme. Für StarCOS 3.0 Karten gelten folgende exemplarische Werte:
  - 18,2 kB Gesamtdatenmenge unterteilt in: 93 Byte MRZ, 16,9 kByte biometrische Daten, 1,17 kByte Verwaltungs- und Sicherheitsdateien
  - Durchschnittliche Lesegeschwindigkeit: 4 bis 5 s für die Gesamtdatenmenge
  - Maximale Lesegeschwindigkeit bei optimiertem Leser: 2 bis 3 s für die GesamtdatenmengeWird die Übertragung mittels Secure Messaging gesichert, verdoppeln sich etwa die Übertragungszeiten.
- **Sicherheitsmechanismen:** Sicherheitsmechanismen sind auf verschiedenen Stufen umsetzbar, die sich im Wesentlichen in der Stärke der Authentisierung und dem Einsatz von Verschlüsselungsverfahren unterscheiden.

## 9 Bewertungsschema

Um die bei BioP II eingesetzten biometrischen Systeme objektiv vergleichend bewerten zu können, ist ein Bewertungsschema mit eindeutigen Kennzahlen notwendig. Nachfolgend werden der allgemeine Aufbau, die Bewertungskriterien und deren Klassifikation sowie die Ergebnisse des für BioP II entwickelten Bewertungsschemas dargestellt.

Auf die detaillierte Darstellung der Ergebnisse zu Überwindungssicherheit, System- und physischer Sicherheit der getesteten Systeme wurde in diesem Bericht verzichtet. Gleichwohl sind diese Kriterien nach der im Folgenden dargestellten Gewichtung in die Bewertung mit eingeflossen.

### 9.1 Aufbau des Bewertungsschemas

Das Bewertungsschema beinhaltet in den einzelnen Zeilen die Bewertungskriterien, die zum Vergleich der biometrischen Systeme herangezogen werden. Die einzelnen Systeme werden innerhalb der Spalten wiedergegeben.

Im Mittelpunkt des Bewertungsschemas steht die Berechnung des System-Index (SI). Ein Index ist in der Statistik ein gewichteter Mittelwert, der aus mehreren Parametern, den Bewertungskriterien, gewonnen wird.

Nicht alle Bewertungskriterien sind für die projektbezogene Beurteilung gleich wichtig. Daher wird jedes Kriterium mit einem Gewicht versehen, welches den Anteil des Kriteriums an der Gesamtnote widerspiegelt.

Der Wertebereich eines jeden Bewertungskriteriums richtet sich nach dem bekannten Schulnotensystem mit Werten von 1 bis 6. Wie die Ergebnisse der Untersuchungen auf diese Noten abgebildet werden, ist in Abschnitt 9.3 dargestellt.

Für jedes Bewertungskriterium wird das Gewicht mit der Note multipliziert. Die Addition der gewichteten Noten ergibt den Indexwert des jeweiligen Systems. Es handelt sich dabei somit um eine Zahl zwischen 1 und 6, die die Gesamtnote wiedergibt.

Im Projekt BioP II interessieren sowohl die Erkennungsleistung der Systeme auf Basis eines ICAO-Bildes als auch auf Basis eines herstellereigenen Templates. Alle Bewertungskriterien, deren Ergebnis sich jeweils für ICAO-Bild und Template unterscheidet, sind in separaten Tabellen zu ICAO (Tabelle 68) und Template (Tabelle 69) zusammengefasst. Innerhalb dieser Tabellen wird eine zusammenfassende Note für diese Kriterien ermittelt, die als ein Wert in das Gesamtbewertungsschema (Tabelle 70) eingeht.

Dieses Vorgehen ermöglicht in einfacher Weise, die Leistung der zu vergleichenden Systeme in Abhängigkeit der verwendeten Referenz (ICAO oder Template) zu bewerten, in dem die Gewichte zwischen ICAO und Template variiert werden.

### 9.2 Bewertungskriterien

Nachfolgend werden die Bewertungskriterien erläutert, die zum Systemvergleich herangezogen werden. Das jeweilige Gewicht gemäß der Relevanz für das Zielszenario kann den Tabellen in Abschnitt 9.4 entnommen werden.

#### 9.2.1 Erkennungsleistung/Kennzahlen

- **FTE:** Die Failure-To-Enrol Rate (FTE) beschreibt den Anteil der Teilnehmer, die nicht an einem biometrischen System enrolt werden können. Die FTE ist in Zusammenhang mit der FRR zu sehen. Die FRR eines biometrischen Systems ist signifikant besser, wenn wiederholte

Merkmalspräsentationen mit geringerer Qualität als nicht-enrolbar klassifiziert werden und somit in die FTE eingehen. Beim realen Einsatz, zum Beispiel bei der Identitätsprüfung an der Grenzkontrolle, geht die FTE direkt in die FRR ein.<sup>49</sup> Daher ist die FTE ein entscheidendes Bewertungskriterium.

- **FRR (FAR=X%)**: Die Erkennungsleistung hat beim Vergleich biometrischer Systeme das höchste Gewicht. Ob dabei FRR oder FAR stärker gewichtet werden, hängt davon ab, ob eher ein akzeptables Sicherheitsniveau oder ein akzeptabler Komfort zu gewährleisten ist. In BioP II wird nach Erkennungsleistungen für verschiedene Sicherheitsniveaus differenziert.
  - FAR=0,001%
  - FAR=0,01%
  - FAR=0,1%
  - FAR=1%

Eine FAR von 0,1% wird als akzeptables Sicherheitsniveau für ein realistisches Einsatzszenario angesehen. Deshalb geht dieses Szenario mit dem größten Gewicht in das Bewertungsschema ein.

In den Punkt FRR(FAR=X%) gehen zu gleichen Teilen die Gesamt-FRR und die gemittelte Einzelbenutzerstatistik zu dem betreffenden Sicherheitsniveau ein.

- **Standardabweichung Einzelbenutzerstatistik**: Eine hohe Standardabweichung besagt, dass ein großer Anteil der Nutzer hohe Rückweisungsraten hat. Die Standardabweichung wird in BioP II in Relation zur Erkennungsleistung schwach gewichtet, da sonst ggf. Algorithmen eine bessere Bewertung erhalten, die zwar eine sehr niedrige Standardabweichung haben, dies aber bei unakzeptabel hohen FRRs. In BioP II besteht ein direkter Zusammenhang zwischen der Standardabweichung und den Benutzbarkeitsproblemen.
- **Größe ICAO-Bild**: Dieses Kriterium bewertet die Größe der als Referenzbasis verwendeten Bilddatei. Die Größe des ICAO-Bildes ist entscheidend hinsichtlich der Speichergröße des zu verwendenden Chips im Personaldokument. Des Weiteren existieren zum Teil Beschränkungen bezüglich maximal übertragbarer Blockgrößen. Nicht zuletzt beeinflusst die Größe des zu übertragenden ICAO-Bildes die Verarbeitungszeit.
- **Größe Template**: Dieses Kriterium bewertet die Größe des als Referenzbasis verwendeten Templates. Grundsätzlich gelten die gleichen Aussagen wie zum Kriterium Größe ICAO-Bild. In der Regel ist die Template-Größe jedoch weniger kritisch, da Templates gewöhnlich signifikant kleiner als ICAO-Bilder sind und im Kontext der Zielanwendung eine nachgeordnete Bedeutung besitzen.

## 9.2.2 Systemergebnisse

- **Mittlere Verarbeitungszeit**: Die mittlere Verarbeitungszeit beschreibt näherungsweise die Zeitdauer, die zur Erkennung einer Person durch das System erforderlich ist. Sie definiert sich aus dem zeitlichen Abstand von Erkennen des aufliegenden Tokens durch das System und dem Vorliegen des Verifikationsergebnisses. Die vollständige Zeit, die eine Person zur Bedienung des Systems benötigt, unterscheidet sich von der Verarbeitungszeit in dem Zeitbedarf für Positionierung und Erkennung des Tokens sowie das visuelle Erfassen des

---

<sup>49</sup> Die Berechnung einer generalisierten FRR, die die FTE berücksichtigt, erfolgt in Abschnitt 7.2.4.

Verifikationsergebnisses durch den Testteilnehmer. Die Ergebnisse zur Bedienzeit wurden im Rahmen der Benutzbarkeitsuntersuchungen ermittelt und sind in Abschnitt 8.1.2.2 dargestellt.

In die Bewertung der mittleren Verarbeitungszeit gehen aufgrund der Vergleichbarkeit ausschließlich Messungen auf Basis der Masterreferenz in Feldtestzyklus I ein.

Bei der Gewichtung der mittleren Verarbeitungszeit wurde berücksichtigt, dass die biometrische Identitätsprüfung im realen Einsatz immer in einen Gesamtprozess (zum Beispiel Einreisekontrolle) eingebettet ist.

- **Systemfehler, Ausfallverhalten, Administrationsaufwand:** Für Systeme, die in Umgebungen mit Publikumsverkehr und daraus resultierenden hohen Betätigungszahlen eingesetzt werden, ist die Gewährleistung eines stabilen Betriebs ein entscheidender Aspekt.
- **Robustheit/Langlebigkeit Sensoren:** Für die geplante Zielanwendung ist ein hoher Durchsatz von Personen an den biometrischen Sensoren zu erwarten. Deshalb spielt dessen Robustheit sowie die Langlebigkeit aufgrund von Folgekosten eine entscheidende Rolle.

### 9.2.3 Benutzerakzeptanz

Die Bewertung der Systeme durch die Benutzer ist eine wichtige Kennzahl für die Akzeptanz, da biometrische Systeme ein kooperatives Verhalten seitens der Benutzer erfordern. Die Benutzungskriterien werden anhand der subjektiven Einschätzungen der Teilnehmer (erfasst mittels Fragebogen) bewertet.

- **Gesamtbeurteilung:** Die subjektive Einschätzung der vier Systeme im direkten Vergleich wird in der Gesamtbewertung wiedergegeben. Die Teilnehmer hatten die Möglichkeit die Systeme in eine Rangliste zwischen 1 (bestes System) und 4 (schlechtestes System) einzuordnen.
- **Schnelligkeit:** Bewertung der subjektiv wahrgenommenen Bedienzeit.
- **Sicherheit:** Bewertung des subjektiven Sicherheitsgefühls, das heißt die Zuverlässigkeit der Erkennung und Wahrscheinlichkeit der Nicht-Überwindung aus der Sicht des Benutzers (im Gegensatz zur objektiven Sicherheit).
- **Einfachheit:** Bewertung der Anzahl und Schwere der Benutzungsprobleme in der Erfahrung oder Erinnerung der Benutzer.
- **Bequemlichkeit:** Bewertung des Aufwands, den die Benutzer bei der Bedienung des Systems leisten müssen.

### 9.2.4 Benutzbarkeit

Die Einschätzung der Benutzbarkeit der Systeme beruht auf objektiven Beobachtungen von Teilnehmern während der Bedienung, einer Analyse des Benutzungsvorgangs sowie Problemnennungen in den Fragebögen.

- **Bedienzeit:** Die Bedienzeit umfasst nicht nur die mittlere Verarbeitungszeit (siehe Abschnitt 9.2.2), sondern die Zeit für den kompletten Bedienungsvorgang seitens der Benutzer.
- **Benutzbarkeitsprobleme:** Dieses Kriterium bewertet die Anzahl und Schwere der beobachteten Benutzungsprobleme klassifiziert nach den Kategorien:
  - Schwierigkeiten der Positionierung (Gewicht:  $\frac{3}{7}$  der Benutzbarkeitsprobleme)

- Mangelndes ergonomisches Design (Gewicht:  $\frac{2}{7}$  der Benutzbarkeitsprobleme)
- Mangelnde Benutzerführung / Rückmeldung (Gewicht:  $\frac{2}{7}$  der Benutzbarkeitsprobleme)

#### 9.2.5 Sicherheit<sup>50</sup>

- **Systemsicherheit:** Dieses Kriterium bewertet die bereitgestellten Systeme hinsichtlich physischer Sicherheit, Sicherheit der Verarbeitungseinheit und Sicherheit der Anwendungssoftware. Da es sich bei den in BioP II eingesetzten Systemen um Prototypen handelt, geht dieses Kriterium nur mit geringem Gewicht in die Bewertung ein.
- **Fake-Resistenz:** Biometrische Systeme müssen insbesondere bei unüberwachtem Betrieb eine ausreichende Resistenz gegen Kopien des betreffenden Merkmals aufweisen. In Abhängigkeit des verwendeten Verfahrens kann aber auch bei überwachtem Betrieb eine ausreichende Fake-Resistenz erforderlich sein. Deshalb stellt die Fake-Resistenz ein wichtiges Bewertungskriterium dar.

#### 9.2.6 Herstellerunterstützung

- **Inbetriebnahme:** Es wird bewertet, inwieweit ein funktionsfähiges System termingerecht bereitgestellt wurde.
- **Support/Service:** Beim Einsatz komplexer technischer Systeme ist der Hersteller-Support ein wichtiges Kriterium.

#### 9.2.7 Nicht zu bewertende Kriterien

- **Einhaltung Normen und Standards:** Dieses Kriterium bewertet die Einhaltung der für biometrische Systeme entwickelten (oder in Arbeit befindlichen) Standards. Für BioP II sind hier insbesondere die Richtlinien und Standards der ICAO bezüglich Charakteristik und Kodierung der Referenzbilder relevant. Die Konformität wurde durch das BKA und secunet bzw. durch Implementierungen von secunet sichergestellt, so dass eine vergleichende Bewertung der Systeme für dieses Kriterium nicht sinnvoll ist.
- **Zertifizierungen und Prüfzeichen:** Dieses Kriterium bewertet das Vorhandensein von Zertifizierungen und Prüfzeichen für die Systeme. Da in BioP II ausschließlich Prototypen zum Einsatz kamen, die als Gesamtsystem in dieser Form vorher noch nicht existierten, erfolgt an dieser Stelle keine Bewertung.
- **Kosten:** Dieses Kriterium bewertet die Kosten für die einzelnen Systeme. Eine Bewertung ist jedoch nicht sinnvoll, da lediglich Prototypen und keine frei verfügbaren Produkte untersucht wurden.

---

<sup>50</sup> Wie bereits weiter oben ausgeführt, wird in diesem Bericht auf die detaillierte Darstellung verzichtet.

### 9.3 Klassifikation der Ergebnisse

In diesem Abschnitt erfolgt die Darstellung der Richtlinien für die Notenvergabe zu den einzelnen Bewertungskriterien. Die Festlegung der Noten erfolgte innerhalb der Projektgruppe und orientiert sich an der Erwartungshaltung an ein biometrisches System im Hinblick auf das Zielszenario Passanwendung.

FTE	Note
$\leq 0,0001\%$	1
$\leq 0,001\%$	2
$\leq 0,01\%$	3
$\leq 0,1\%$	4
$\leq 1\%$	5
$> 1\%$	6

**Tabelle 48: Notenklassifikation zur FTE**

FRR	Note
$\leq 1\%$	1
$\leq 2\%$	2
$\leq 4\%$	3
$\leq 8\%$	4
$\leq 16\%$	5
$> 16\%$	6

**Tabelle 49: Notenklassifikation zur FRR**

Standardabweichung in der Einzelbenutzerstatistik	Note
$\leq 1\%$	1
$\leq 2\%$	2
$\leq 4\%$	3
$\leq 8\%$	4
$\leq 16\%$	5
$> 16\%$	6

**Tabelle 50: Notenklassifikation zur Standardabweichung in der Einzelbenutzerstatistik**

ICAO-Bild-Größe	Note
≤ 4 kByte	1
≤ 8 kByte	2
≤ 16 kByte	3
≤ 32 kByte	4
≤ 64 kByte	5
> 64 kByte	6

**Tabelle 51: Notenklassifikation zur ICAO-Bild-Größe**

Template-Größe	Note
≤ 0,5 kByte	1
≤ 1 kByte	2
≤ 2 kByte	3
≤ 4 kByte	4
≤ 8 kByte	5
> 8 kByte	6

**Tabelle 52: Notenklassifikation zur Template-Größe**

Mittlere Verarbeitungszeit	Note
≤ 2 s	1
≤ 4 s	2
≤ 6 s	3
≤ 8 s	4
≤ 10 s	5
> 10 s	6

**Tabelle 53: Notenklassifikation zur mittleren Verarbeitungszeit**

Für die Gruppierung der Systemfehler und deren Behebung gelten folgende Definitionen.

- Gruppierung der Fehler
  - **Marginale Fehler:** Das Problem hat keine Auswirkung auf die Betriebsbereitschaft und das Management des Systems.  
Beispiel: Performance-Probleme der Cognitec-Systeme
  - **Kleine Fehler:** Die Betriebsbereitschaft und das Management des Systems sind nur in geringem Umfang beeinträchtigt.

Beispiel: Wackelkontakt an der LED für das Benutzerfeedback

- **Fehler:** Die Betriebsbereitschaft des Systems ist nicht gegeben. Es handelt sich um einen Totalausfall des Systems.

Beispiel: Abbruch der Verifikationssoftware

- Behebung des Fehlers:
  - **Einfache Maßnahmen:** Durch eine einfache Fehlerdiagnose konnte die Ursache ermittelt werden. Seitens des Herstellers wurde zeitnah (innerhalb 1-2 Tagen) ein Update (bzw. eine Ersatzhardware) bereitgestellt, das die Fehlerursache behoben hat.
  - **Mittlerer Aufwand:** Es waren mehrere Fehlerdiagnosen notwendig, um die Ursache zu ermitteln. Seitens des Herstellers wurden mehrere Updates bereitgestellt. Die Fehlerursache konnte zeitnah behoben werden.
  - **Hoher Aufwand:** Es wurden mehrere komplexe und zeitaufwendige Fehlerdiagnosen durchgeführt. Seitens des Herstellers wurden mehrere Updates bereitgestellt. Die Behebung der Fehlerursache erfolgte nicht zeitnah.
  - **Fehler konnte nicht behoben werden:** Der Fehler konnte während der gesamten Feldtestlaufzeit nicht behoben werden. Es besteht jedoch Aussicht, den Fehler durch Einsatz geeigneter Diagnosemaßnahmen zu eruieren und beheben zu können.
  - **Zweckmäßiger Einsatz verhindert:** Der Fehler konnte während der gesamten Feldtestlaufzeit nicht behoben werden. Nach derzeitigem Stand der Technik besteht keine Aussicht, diesen Fehler durch geeignete Diagnosemaßnahmen zu ermitteln und zu beheben. Ein zweckmäßiger Einsatz des Systems ist daher nicht möglich.

Systemfehler	Note
Keine oder nur marginale Fehler während des Feldtests	1
Kleine Fehler während des Feldtests, die durch einfache Maßnahmen behoben werden konnten	2
Fehler während des Feldtests, die durch Maßnahmen mittleren Aufwandes behoben werden konnten	3
Fehler während des Feldtests, die durch Maßnahmen hohen Aufwands behoben werden konnten	4
Fehler während des Feldtests, die nicht behoben werden konnten	5
Fehler während des Feldtests, die einen zweckgemäßen Einsatz verhinderten	6

**Tabelle 54: Notenklassifikation zu Systemfehlern**

Ausfallverhalten	Note
Weniger als 16 Totalausfälle während des Feldtests	1
Weniger als 32 Totalausfälle während des Feldtests	2
Weniger als 48 Totalausfälle während des Feldtests	3
Weniger als 64 Totalausfälle während des Feldtests	4
Weniger als 128 Totalausfälle während des Feldtests	5
128 oder mehr Totalausfälle während des Feldtests	6

**Tabelle 55: Notenklassifikation zum Ausfallverhalten**

Kriterien für die Bewertung des Administrationsaufwands:

- Betreuung im Betrieb
- Änderung der Konfiguration
- Systemtransfer
- Tätigkeiten bei Problemen
- Enrolment
- Diagnoseprogramme

Administrationsaufwand	Note
Einhaltung aller Kriterien	1
Ein Kriterium wird nicht erfüllt	2
Zwei Kriterien werden nicht erfüllt	3
Drei Kriterien werden nicht erfüllt	4
Vier Kriterien werden nicht erfüllt	5
Mehr als vier Kriterien werden nicht erfüllt	6

**Tabelle 56: Notenklassifikation zum Administrationsaufwand**

Langzeitstabilität der Sensoren		Note
Änderung Bildqualität	Physischer Verschleiß	
Keine Qualitätseinbußen	Kein Verschleiß	1
Geringe Qualitätseinbußen	Geringer Verschleiß	2
Deutliche Qualitätseinbußen	Deutlicher Verschleiß	3
Erhebliche Qualitätseinbußen	Erheblicher Verschleiß	4
Große Qualitätseinbußen	Großer Verschleiß	5
Sehr große Qualitätseinbußen	Sehr großer Verschleiß	6

**Tabelle 57: Notenklassifikation zu Robustheit/Langlebigkeit der Sensoren<sup>51</sup>**

Benutzerakzeptanz – Gesamtbeurteilung	Note
Rang im Mittel zwischen 1 und 1,5	1
Rang im Mittel zwischen 1,5 und 2	2
Rang im Mittel zwischen 2 und 2,5	3
Rang im Mittel zwischen 2,5 und 3	4
Rang im Mittel zwischen 3 und 3,5	5
Rang im Mittel zwischen 3,5 und 4	6

**Tabelle 58: Notenklassifikation zur Gesamtbeurteilung**

Benutzerakzeptanz – Benutzungskriterien	Note
Bewertung des Systems im Mittelwert < 2	1
Bewertung des Systems im Mittelwert < 3	2
Bewertung des Systems im Mittelwert < 4	3
Bewertung des Systems im Mittelwert < 5	4
Bewertung des Systems im Mittelwert < 6	5
Bewertung des Systems im Mittelwert = 6	6

**Tabelle 59: Notenklassifikation zu Benutzungskriterien**

<sup>51</sup> Klassifikation nach erwartetem bzw. nachweisbarem Einfluss auf die Erkennungsleistung

Bedienzeit	Note
≤ 8 s	1
≤ 10 s	2
≤ 12 s	3
≤ 14 s	4
≤ 16 s	5
> 16 s	6

**Tabelle 60: Notenklassifikation zur Bedienzeit**

Benutzbarkeitsprobleme	Note
≤ 1% Probleme	1
≤ 2% Probleme	2
≤ 3% Probleme	3
≤ 4% Probleme	4
≤ 5% Probleme	5
> 5% Probleme	6

**Tabelle 61: Notenklassifikation zu Benutzbarkeitsproblemen**

Die Note für die Systemsicherheit setzt sich aus drei Teilnoten zusammen, die zu gleichen Teilen in die Gesamtnote eingehen:

- Physische Sicherheit
- Sicherheit der Verarbeitungseinheit
- Sicherheit der Anwendungssoftware

Physische Sicherheit	Note
Kein Kriterium kritisch	1
-	2
-	3
Ein Kriterium kritisch	4
Zwei Kriterien kritisch	5
Alle drei Kriterien kritisch	6

**Tabelle 62: Notenklassifikation zur physischen Sicherheit**

Sicherheit der Verarbeitungseinheit	Note
Kein Kriterium kritisch	1
-	2
Ein Kriterium kritisch	3
Zwei Kriterien kritisch	4
Drei Kriterien kritisch	5
Vier Kriterien kritisch	6

**Tabelle 63: Notenklassifikation zur Sicherheit der Verarbeitungseinheit**

Sicherheit der Anwendungssoftware	Note
Kein Kriterium kritisch	1
Ein Kriterium kritisch	2
Zwei Kriterien kritisch	3
Drei Kriterien kritisch	4
Vier Kriterien kritisch	5
Fünf oder sechs Kriterien kritisch	6

**Tabelle 64: Notenklassifikation zur Sicherheit der Anwendungssoftware**

Fake-Resistenz	Note
Es gibt keinen Erfolg versprechenden Ansatzpunkt für Überwindungsversuche	1
Überwindung mit mittlerem Aufwand sowie Spezialwissen und/oder speziellen Hilfsmitteln nicht erfolgreich; es gibt jedoch einen Erfolg versprechenden Ansatzpunkt	2
Überwindung mit mittlerem Aufwand sowie Spezialwissen und/oder speziellen Hilfsmitteln erfolgreich	3
Überwindung mit mittlerem Aufwand erfolgreich (mit Zugriff auf das Merkmal eines Berechtigten)	4
Überwindung mit mittlerem Aufwand erfolgreich (ohne Zugriff auf das Merkmal eines Berechtigten)	5
Überwindung mit niedrigem Aufwand erfolgreich	6

**Tabelle 65: Notenklassifikation zur Fake-Resistenz**

Inbetriebnahme	Note
Termingerechte Bereitstellung eines voll funktionsfähigen Systems	1
Termingerechte Bereitstellung des Systems, aber nachträgliche Anpassungen erforderlich	2
Termingerechte Bereitstellung des Systems und wesentliche nachträgliche Anpassungen erforderlich oder verspätete Bereitstellung eines voll funktionsfähigen Systems	3
Verspätete Bereitstellung des Systems und nachträgliche Anpassungen erforderlich	4
Verspätete Bereitstellung des Systems und wesentliche nachträgliche Anpassungen erforderlich	5
Keine Bereitstellung eines geeigneten Systems	6

**Tabelle 66: Notenklassifikation zur Inbetriebnahme**

Kriterien für die Bewertung von Support / Service:

- Reaktion innerhalb 24h
- Problemlösung innerhalb 24 h
- Alle auftretenden Fehlerfälle wurden behoben
- Benennung eines kompetenten Ansprechpartners
- Bereitstellung einer geeigneten Dokumentation

Support / Service	Note
Einhaltung aller Kriterien	1
Ein Kriterium wird nicht erfüllt	2
Zwei Kriterien werden nicht erfüllt	3
Drei Kriterien werden nicht erfüllt	4
Vier Kriterien werden nicht erfüllt	5
Alle fünf Kriterien werden nicht erfüllt	6

**Tabelle 67: Notenklassifikation zu Support / Service**

#### **9.4 Systembewertung**

In den folgenden Tabellen sind die Schemata zur Bewertung der ICAO- und Template-spezifischen Bewertungskriterien sowie das Gesamtbewertungsschema dargestellt.

Kriterium	Gew.	Cognitec		BDr/NEC		Dermalog		SDI	
		Note	Ergeb.	Note	Ergeb.	Note	Ergeb.	Note	Ergeb.
<b>Gesamt-FRR</b>									
FRR(FAR=0,001%)	8,0%	6	0,48	4	0,32	4	0,32	5	0,4
FRR(FAR=0,01%)	8,0%	6	0,48	4	0,32	4	0,32	5	0,4
FRR(FAR=0,1%)	22,0%	4	0,88	3	0,66	3	0,66	4	0,88
FRR(FAR=1%)	5,0%	3	0,15	3	0,15	2	0,1	4	0,2
<b>EBS gemittelt</b>									
FRR(FAR=0,001%)	8,0%	6	0,48	4	0,32	4	0,32	6	0,48
FRR(FAR=0,01%)	8,0%	6	0,48	4	0,32	4	0,32	5	0,4
FRR(FAR=0,1%)	22,0%	4	0,88	4	0,88	4	0,88	5	1,1
FRR(FAR=1%)	5,0%	4	0,2	4	0,2	3	0,15	5	0,25
<b>Sonstiges</b>									
FTE	10,0%	1	0,1	5	0,5	1	0,1	1	0,1
Std.Abweichung EBS	2,0%	5	0,1	5	0,1	5	0,1	6	0,12
Mittlere ICAO-Bild-Größe	2,0%	3	0,06	3	0,06	3	0,06	5	0,1
<b>Summe</b>	<b>100,0%</b>		<b>4,29</b>		<b>3,83</b>		<b>3,33</b>		<b>4,43</b>

Tabelle 68: ICAO-spezifische Bewertungskriterien

Kriterium	Gew.	Cognitec		BDr/NEC		Dermalog		SDI	
		Note	Ergeb.	Note	Ergeb.	Note	Ergeb.	Note	Ergeb.
<b>Gesamt-FRR</b>									
FRR(FAR=0,001%)	8,0%	6	0,48	3	0,24	3	0,24	4	0,32
FRR(FAR=0,01%)	8,0%	4	0,32	3	0,24	3	0,24	4	0,32
FRR(FAR=0,1%)	22,0%	3	0,66	3	0,66	2	0,44	4	0,88
FRR(FAR=1%)	5,0%	3	0,15	3	0,15	2	0,1	4	0,2
<b>EBS gemittelt</b>									
FRR(FAR=0,001%)	8,0%	6	0,48	4	0,32	4	0,32	5	0,4
FRR(FAR=0,01%)	8,0%	4	0,32	4	0,32	3	0,24	5	0,4
FRR(FAR=0,1%)	22,0%	4	0,88	4	0,88	3	0,66	5	1,1
FRR(FAR=1%)	5,0%	3	0,15	4	0,2	3	0,15	5	0,25
<b>Sonstiges</b>									
FTE	10,0%	1	0,1	5	0,5	5	0,5	5	0,5
Std.Abweichung EBS	2,0%	5	0,1	5	0,1	5	0,1	6	0,12
Mittlere Template-Größe	2,0%	6	0,12	1	0,02	1	0,02	1	0,02
<b>Summe</b>	<b>100,0%</b>		<b>3,76</b>		<b>3,63</b>		<b>3,01</b>		<b>4,51</b>

Tabelle 69: Template-spezifische Bewertungskriterien

Kriterium	Gew.	Cognitec		BDr/NEC		Dermalog		SDI	
		Note	Ergeb.	Note	Ergeb.	Note	Ergeb.	Note	Ergeb.
<b>Erkennungsleistung / Kennzahlen</b>									
ICAO	48,0%	4,29	2,059	3,830	1,838	3,330	1,598	4,430	2,126
Template	12,0%	3,76	0,451	3,630	0,436	3,010	0,361	4,510	0,541
<b>Systemergebnisse</b>									
Mittlere Verarbeitungszeit	0,5%	4	0,02	4	0,02	3	0,015	3	0,015
Systemfehler	1,0%	5	0,05	3	0,03	4	0,04	4	0,04
Ausfallverhalten	1,0%	5	0,05	3	0,03	3	0,03	3	0,03
Administrationsaufwand	0,5%	5	0,025	2	0,01	3	0,015	2	0,01
Robustheit/Langlebigkeit Sensoren	2,0%	1	0,02	2	0,04	1	0,02	1	0,02
<b>Benutzerakzeptanz</b>									
Gesamtbeurteilung	3,0%	4	0,12	3	0,09	3	0,09	2	0,06
Sicherheit	3,0%	2	0,06	2	0,06	2	0,06	2	0,06
Schnelligkeit	3,0%	3	0,09	2	0,06	2	0,06	2	0,06
Einfachheit	3,0%	1	0,03	1	0,03	1	0,03	1	0,03
Bequemlichkeit	3,0%	1	0,03	1	0,03	1	0,03	1	0,03
<b>Benutzbarkeit</b>									
Bedienzeit	3,0%	3	0,09	4	0,12	4	0,12	4	0,12
Benutzbarkeitsprobleme	7,0%	2	0,14	1	0,07	2	0,14	4	0,28
<b>Sicherheit</b>									
Systemsicherheit	1,0%	2	0,02	6	0,06	4,7	0,047	5,3	0,053
Fake-Resistenz	4,0%	4	0,16	4	0,16	4	0,16	2	0,08
<b>Sonstiges</b>									
Inbetriebnahme	2,0%	3	0,06	2	0,04	3	0,06	3	0,06
Support/Service	3,0%	3	0,09	2	0,06	2	0,06	2	0,06
<b>Summe</b>	<b>100,0%</b>		<b>3,57</b>		<b>3,18</b>		<b>2,94</b>		<b>3,68</b>

Tabelle 70: Gesamtbewertungsschema

## 10 Zusammenfassung und Interpretation der Ergebnisse

Ziel des Projekts BioP II war die Untersuchung der Leistungsfähigkeit von biometrischen Verifikationssystemen im Hinblick auf den Einsatz dieser Technologie in Personaldokumenten. Dabei standen die folgenden Fragen im Vordergrund:

- Welche Erkennungsleistung bieten die Verfahren Gesichts-, Fingerabdruck- und Iriserkennung?
- Können die Verfahren mit Bilddateien gemäß ICAO als Referenz statt mit proprietären herstellereigenen Referenz-Templates zufrieden stellende Ergebnisse erzielen?
- Welche Praxistauglichkeit bieten die drei Verfahren für eine große Testgruppe? Wie ist es um Benutzbarkeit und Akzeptanz bestellt?
- Welches der getesteten Verfahren eignet sich unter welchen Bedingungen am ehesten für die geplante Verwendung in Personaldokumenten?<sup>52</sup>

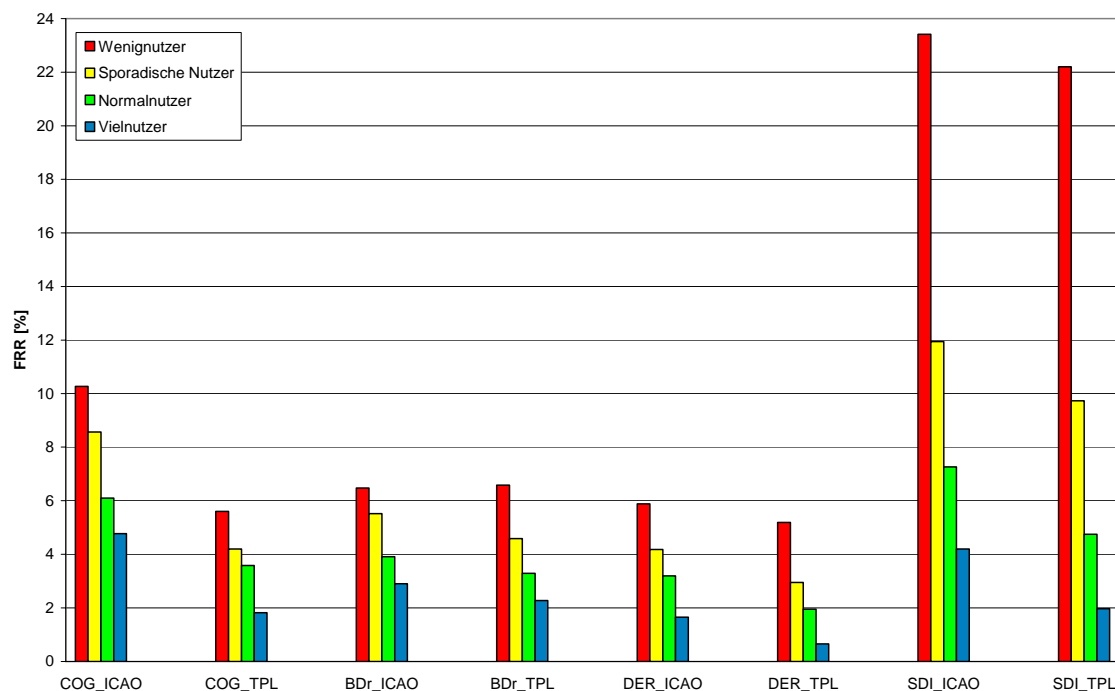
Die Beantwortung dieser Fragestellungen kann nur auf Basis konkreter Implementierungen biometrischer Systeme erfolgen. Neben der Betrachtung der genannten Fragestellungen erfolgt daher auch eine allgemeine Einschätzung der Kompetenz von Anbietern biometrischer Systeme. Abschließend werden auf Basis der gewonnenen Ergebnisse Empfehlungen zur Einführung biometrischer Verfahren zur Verwendung mit Personaldokumenten aufgezeigt.

### 10.1 Erkennungsleistung der biometrischen Verfahren

Die Erkennungsleistung der in BioP II eingesetzten biometrischen Systeme war hochgradig nutzerklassenspezifisch. Personen, die viele Betätigungen durchgeführt haben und somit entsprechend trainiert waren, erzielten vergleichsweise deutlich geringere Falschrückweisungsrate als ungeübte Nutzer. Zur Veranschaulichung ist in Abbildung 105 die Falschrückweisungsrate (FRR) der eingesetzten Systeme jeweils für die Referenzen ICAO-Bild und Template bei einem vergleichbaren Sicherheitsniveau (FAR=0,1%) dargestellt.

---

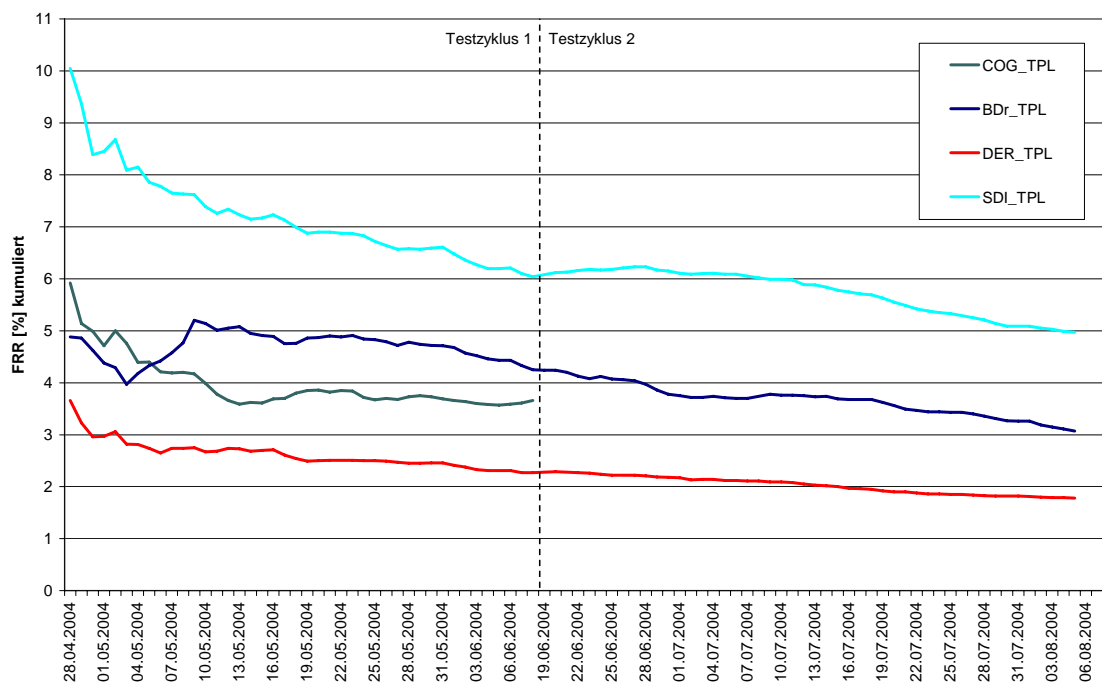
<sup>52</sup> Diese Fragestellung wurde teilweise von den nationalen und europäischen Entwicklungen während der bereits laufenden Untersuchungen überholt, da zwischenzeitlich die Entscheidungen, wie im Vorwort benannt, getroffen wurden.



**Abbildung 105: Nutzerklassen-spezifische Erkennungsleistung (FAR=0,1%)**

Personen, die nur wenige Betätigungen durchführten, verursachten höhere Rückweisungsdaten als geübte Nutzer. Auch im Hinblick auf die unterschiedlichen Nutzerklassen muss also ein biometrisches System anwendungsspezifisch beurteilt werden. Während etwa bei einem System, das typischerweise sehr häufig und von Vielnutzern bedient wird, durch den schnell erzielbaren Trainingseffekt vermutlich niedrige FRRs erreichbar sind, ist für eine Anwendung mit seltenen Nutzern aufgrund deren Ungeübtheit mit höheren Rückweisungsdaten zu rechnen. Auch für diese Nutzergruppe ist allerdings bei zunehmender Nutzung biometrischer Systeme allgemein mit verbesserten Rückweisungsdaten zu rechnen.

Der Trainingseffekt ist für die verschiedenen Verfahren durchaus unterschiedlich. Dies wird durch die in Abbildung 106 dargestellte Entwicklung der Falschrückweisungsrate im zeitlichen Verlauf deutlich. Während sich bei dem Gesichtserkennungssystem nach einer bestimmten Zeit die FRR stabilisierte, war auch zum Ende des Feldtests bei beiden Fingerabdruckerkennungssystemen weiterhin ein schwacher und beim Iriserkennungssystem ein deutlicher Rückgang der FRR zu verzeichnen. Insbesondere für das Iriserkennungssystem ist daher ein erhebliches Verbesserungspotenzial durch intensives Training der Nutzer zu erwarten. Für das Zielszenario Einsatz in Personaldokumenten ist dies jedoch weniger relevant, da hier die Benutzer in unterschiedlicher Frequenz mit der biometrischen Anwendung konfrontiert werden. Die Schwierigkeiten beim Iriserkennungssystem waren im Wesentlichen auf die verwendete Kameraeinheit zurückzuführen. Mit einer anderen Sensoreinheit, die auch für ungeübte Nutzer eine einfachere Merkmalspräsentation ermöglicht und über ein geeignetes Benutzerfeedback verfügt, ist eine deutliche Verbesserung der Erkennungsleistung zu erwarten. Darüber hinaus wird der zunehmende Einsatz biometrischer Systeme im Alltag zu einer Gewöhnung und Übung der Nutzer führen, die diesen Effekt der verbesserten Erkennungsleistung noch verstärken werden.



**Abbildung 106: Erkennungsleistung im zeitlichen Verlauf am Beispiel Template (FAR=0,1%)**

Mit den getesteten Verfahren lassen sich für unterschiedliche Sicherheitsniveaus in Abhängigkeit der verwendeten Referenzbasis sowie der Nutzergruppe die in Tabelle 71 dargestellten Bereiche von Erkennungsleistungen erreichen.

	FAR=0,001%	FAR=0,01%	FAR=0,1%	FAR=1%
Gesichtserkennung	15% – 50%	3% – 23%	2% – 10%	1% – 7%
Fingerabdruckerkennung	2% – 10%	1% – 8%	1% – 7%	0% – 5%
Iriserkennung	3% – 26%	2% – 25%	2% – 23%	2% – 23%

**Tabelle 71: Bereiche erzielbarer FRRs**

Die durchgeführten Untersuchungen haben gezeigt, dass bei einer Erhöhung des Sicherheitsniveaus die Leistungsfähigkeit des Gesichtserkennungssystems stärker abnimmt als bei den untersuchten Fingerabdruck- und Iriserkennungssystemen.

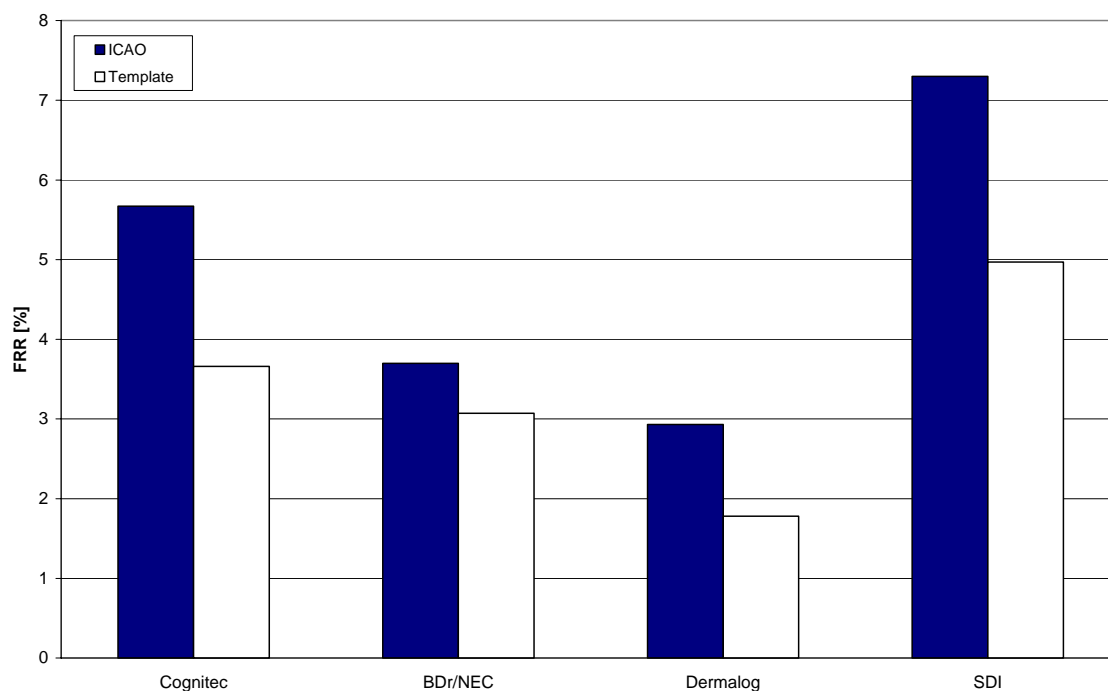
Bei der Betrachtung der in BioP II ermittelten Falschrückweisungsrate muss berücksichtigt werden, dass mit der biometrischen Verifikation keine echte Anwendung (wie beispielsweise eine Zutrittskontrolle) verknüpft war, es sich also um einen reinen Testbetrieb handelte. In realen Einsatzumgebungen ist aufgrund der zu erwartenden erhöhten Kooperationsbereitschaft mit geringeren Rückweisungsrate zu rechnen, da hier in aller Regel davon auszugehen ist, dass die Nutzer am Erfolg der Verifikation ein unmittelbares Interesse haben. Des Weiteren werden biometrische Systeme in der Praxis in aller Regel in der Weise betrieben, dass die Möglichkeit mehrerer direkt aufeinander folgender Versuche für die Identitätsprüfung besteht. Auf diese Einstellung wurde bei BioP II aus Gründen der Testmethodik verzichtet und nur die jeweils erste Betätigung eines Versuchs ausgewertet. Durch die Kumulation mehrerer Versuche ist eine Verbesserung der FRR zu erwarten.

Hinsichtlich möglicher Verbesserungen der Erkennungsleistung biometrischer Systeme ist die Weiterentwicklung der Algorithmen und Sensoren nur ein Aspekt. Wesentlich größeres Potenzial liegt

in einer Optimierung der Ergonomie und des Benutzerfeedbacks sowie des Benutzertrainings, um den Nutzern die Präsentation des betreffenden biometrischen Merkmals in der erforderlichen Qualität zu erleichtern.

## 10.2 Einsetzbarkeit ICAO-Bilddateien als Referenz

Die biometrischen Systeme wurden auch mit Bilddateien gemäß den ICAO-Richtlinien als Referenz betrieben.



**Abbildung 107: Vergleich Erkennungsleistung ICAO-Bild vs. Template (FAR=0,1%)**

Abbildung 107 zeigt, dass beim Sicherheitsniveau FAR=0,1% ICAO-Bilder als Referenz zu einer um 1% bis zu etwa 2,5% höheren FRR führen als Referenz-Templates. Während bei den Fingerabdruckerkennungssystemen auch bei höheren Sicherheitsniveaus der Unterschied in der Erkennungsleistung etwa in diesem Bereich liegt und bei dem Iriserkennungssystem für ICAO-Bilder eine um bis zu 5% höhere FRR auftritt, ist für das Gesichtserkennungssystem ein anderes Verhalten zu beobachten. Schon bei FAR=0,01% ist die FRR für das ICAO-Bild um ca. 10% höher als die FRR für das Template.

Die Erkennungsleistung bei der Verwendung von ICAO-Referenzbildern war also niedriger als bei Verwendung von Referenz-Templates, die beim System-Enrolment aufgenommen wurden. Dies kann damit begründet werden, dass die zurzeit verfügbaren biometrischen Systeme in der Regel auf die herstellereigenen proprietären Templates trainiert sind. Die Erkennungsleistung mit ICAO-Referenzbildern bewegte sich jedoch immer noch in einer akzeptablen Größenordnung (vergleiche Abbildung 2). Ein Einsatz biometrischer Verfahren mit ICAO-Bildern als Referenz ist daher bereits heute mit zufriedenstellenden Ergebnissen möglich. Weitere Verbesserungen sind durch die Entwicklungen in der Standardisierung biometrischer Verfahren zu erwarten. Bereits heute vorhandene Optimierungspotentiale der Algorithmen bei der Verarbeitung entsprechender Bilddaten können so weiter ausgebaut werden.

Beim Vergleich der Erkennungsleistung von ICAO-Bild und Template ist ausserdem zu beachten, dass in BioP II bei den Fingerabdruck- und Iriserkennungssystemen eine Benachteiligung der ICAO-Referenz gegeben war, da das Template als Masterreferenz konfiguriert war und damit eine hierfür optimierte Aufnahme des Live-Bildes erfolgte. Auch für das Gesichtserkennungssystem ist ein Vergleich der beiden Referenzbasen nur eingeschränkt möglich, da die Ergebnisse aus verschiedenen Auswertungszeiträumen resultieren, für die unterschiedliche Umgebungsbedingungen und Systemkonfigurationen gegeben waren.

### **10.3 Benutzbarkeit und Akzeptanz**

#### **10.3.1 Benutzbarkeit**

Es konnte festgestellt werden, dass bei allen eingesetzten Systemen das Design und das Benutzerfeedback verbesserungsfähig sind. Dabei ist allerdings auch zu berücksichtigen, dass den Herstellern nur ein stark begrenzter Zeitraum für die Konstruktion der Systeme unter Berücksichtigung der Vorgaben des Auftraggebers zur Verfügung stand. Die Gestaltung der Systeme sollte unter anderem so erfolgen, dass der Benutzer bei einer möglichst optimalen Merkmalspräsentation unterstützt wird. Während bei Vielnutzern eine Benutzerführung kaum notwendig ist (hinsichtlich der dann unter Umständen längeren Bedienzeit ggf. sogar kontraproduktiv für die Akzeptanz), muss bei Wenignutzern auf diesen Aspekt sehr viel Wert gelegt werden. Bei dieser Gruppe ist eine ggf. längere Bedienzeit weniger relevant, wenn die Bedienung trotz seltener Benutzung problemlos und einfach erfolgen kann.

Abgesehen von durch das Systemdesign (zum Beispiel unklares Feedback, mangelnde Performance) verursachten Problemen waren bei dem Gesichtserkennungssystem keine grundsätzlichen Schwierigkeiten bei der Präsentation des Merkmals zu beobachten. Bei den Fingerabdruckerkennungssystemen traten teilweise Probleme bei der Merkmalspräsentation hinsichtlich Auflagendruck und Positionierung auf. Da das in BioP II eingesetzte Iriserkennungssystem eine sehr genaue Positionierung des Auges vor der Erfassungseinheit erforderte, führte dies insbesondere bei ungeübten Personen zu Schwierigkeiten bei der Erkennung.

#### **10.3.2 Akzeptanz**

Die Testteilnehmer von BioP II standen dem Einsatz von Biometrie grundsätzlich sehr positiv gegenüber. Dies drückte sich in einer Zustimmung zum Einsatz biometrischer Verfahren für verschiedenste Anwendungszwecke aus. Gleichzeitig wurde großer Informationsbedarf u.a. hinsichtlich der Aspekte der Erfassung (Gesundheit, Hygiene) und der Verarbeitung bzw. Speicherung (Datenschutz) biometrischer Daten festgestellt. Bei der ermittelten hohen Akzeptanz ist zwar zu berücksichtigen, dass es sich bei BioP II um freiwillige Testteilnehmer handelte, bei denen eine positive Grundeinstellung vermutet werden kann. Jedoch haben sich auch solche Personen für die Teilnahme am Feldtest entschieden, die nach ihren eigenen Angaben gegenüber biometrischen Verfahren grundsätzlich kritisch eingestellt waren.

### **10.4 Eignung der getesteten Verfahren**

#### **10.4.1 Gesichtserkennung**

Die **Gesichtserkennung** ist mit dem hier verwendeten System einfach benutzbar. Die Präsentation des Merkmals stellt für die Nutzer in der Regel kein Problem dar, so dass kaum Training erforderlich ist. Insbesondere für Anwendungen mit Wenignutzern ist dies ein Vorteil.

Im Hinblick auf den Einsatz in Personaldokumenten ist von Vorteil, dass bereits in der herkömmlichen Form Lichtbilder verwendet werden. Für eine digitale Verwendung muss zusätzlich sichergestellt werden, dass die verwendeten Bilder eine für die Gesichtserkennung geeignete Qualität aufweisen.

BioP II hat die Erfahrungen aus BioP I bestätigt und damit die Abhängigkeit der Gesichtserkennung von den Lichtbedingungen in der Umgebung. Bereits kleine Änderungen der Beleuchtungsverhältnisse haben bei dem hier eingesetzten System zu starken Änderungen der Erkennungsleistung geführt. Mit alternativen Erfassungseinheiten können vermutlich andere Ergebnisse erzielt werden. Insbesondere die Verwendung von ICAO-Bildern als Referenz erfordert eine homogene Ausleuchtung des Gesichts bei der Aufnahme des Live-Bildes. Bei sehr unterschiedlichen Bedingungen an potenziellen Einsatzorten ist daher mit verschiedenen Resultaten in der Erkennungsleistung zu rechnen. Auch beim Einsatz von Templates muss für vergleichbare Lichtbedingungen bei Enrolment und Verifikation gesorgt werden.

Schließlich kann Gesichtserkennung grundsätzlich von allen Ausweisinhabern genutzt werden, es können alle Gesichtsbilder erfasst werden. Ausnahmen bilden Personen, für die sich etwa aus religiösen bzw. kulturellen Gründen eine Aufnahme des Gesichts verbietet, oder aus anderen Gründen wesentliche Teile des Gesichts verdeckt sind.

#### 10.4.2 Fingerabdruckerkennung

Die **Fingerabdruckerkennung** liefert mit den beiden hier untersuchten Systemen im Vergleich zu den beiden weiteren Verfahren die beste Erkennungsleistung. Beide Fingerabdruckerkennungssysteme können bei den hier betrachteten Sicherheitsniveaus die guten Erkennungsraten der beiden anderen Verfahren noch überbieten. Die Ergebnisse zeigen damit, dass die Fingerabdruckerkennung auch für den Einsatz bei sehr hohen Sicherheitsniveaus geeignet ist.

Bei beiden Fingerabdruckerkennungssystemen kamen im Feldtest optische Sensoren zum Einsatz. Trotz der hohen Frequentierung (bis zu 18.000 Betätigungen pro Sensor) wiesen die Sensoren eine sehr hohe Stabilität auf. Es konnte bei mehrfach täglicher Reinigung weder eine nennenswerte Verschlechterung der Bildqualität beobachtet werden, noch gab es Gebrauchs- oder Verschleißspuren.

Aufgrund z.B. schwacher Merkmalsausprägung oder körperlichen Einschränkungen, die eine geeignete Präsentation des Merkmals verhindert, kann ein gewisser Anteil potenzieller Nutzer die Fingerabdruckerkennung nicht benutzen. Der Anteil dieser nicht-enrolbaren Personen könnte allerdings reduziert werden, indem beispielsweise der Daumen verwendet wird, der in der Regel über eine höhere Anzahl von auswertbaren Merkmalen verfügt als andere Finger.

#### 10.4.3 Iriserkennung

Die Erkennungsleistung des hier untersuchten **Iriserkennungssystems** würde den Einsatz für hohe Sicherheitsniveaus ermöglichen. Es ist festzuhalten, dass die Falschakzeptanzrate (FAR) für die fest implementierte Toleranzschwelle höher ist als von den Herstellern regelmäßig angegeben. Geeigneter erscheint daher, den fest implementierten Schwellwert nicht als Dogma anzusehen, sondern auch hier Konfigurationsmöglichkeiten hinsichtlich verschiedener Sicherheitsniveaus zu schaffen.




Festzustellen waren bei dem hier getesteten Iriserkennungssystem die Bedienungsschwierigkeiten für ungeübte Benutzer. Da das zum Einsatz kommende Kamerasystem eine genaue Positionierung des Auges erfordert, führte dies bei Personen, die das System nur selten benutzten, gehäuft zu Problemen, die sich für in einer FRR von über 20% in der Gruppe der Wenignutzer widerspiegelten. Dieser Effekt war jedoch auch bei einzelnen Personen festzustellen, die die Iriserkennungssysteme regelmäßig benutzten. Mit einer anderen Sensoreinheit, die auch für ungeübte Nutzer eine einfache Merkmalspräsentation ermöglicht und über eine geeignete Benutzerführung verfügt, ist eine deutliche Verbesserung der Erkennungsleistung zu erwarten. Darüber hinaus wird der zunehmende Einsatz

biometrischer Systeme im Alltag zu einer Gewöhnung und Übung der Nutzer führen, die zu einer weiter verbesserten Erkennungsleistung führen werden.

Auch bei der Iriserkennung ist grundsätzlich zu berücksichtigen, dass ein gewisser Anteil von Personen diese nicht nutzen können, weil diesen z.B. aufgrund starker Sehschwäche oder anderer körperlicher Einschränkungen eine geeignete Präsentation der Iris nicht möglich ist.

#### 10.4.4 Vergleichende Bewertung

Aufgrund der im Rahmen der Untersuchungen erreichten Ergebnisse ist eine Bewertung der getesteten Systeme anhand verschiedener Kriterien möglich. Diese Bewertung ist, da anhand konkreter biometrischer Systeme in einem bestimmten Anwendungsszenario getroffen, nicht ohne weiteres verallgemeinerungsfähig. Die hier erzielten Ergebnisse zeigen jedoch Tendenzen der aktuellen Leistungsfähigkeit der betroffenen Verfahren auf und können deshalb zur Beantwortung wesentlicher Fragen für den hier relevanten Einsatzzweck herangezogen werden. Ausgehend vom angestrebten Einsatzzweck sind die Bewertungskriterien unterschiedlich zu gewichten. Für das Szenario Identitätskontrolle in Zusammenhang mit Personaldokumenten ergibt sich anhand der hier untersuchten Systeme ein Vorsprung der Fingerabdruckerkennung vor den beiden anderen Technologien, von denen die Gesichtserkennung etwas besser abschneidet als die Iriserkennung.

 <p>1</p>	<ul style="list-style-type: none"> <li>+ Beste Erkennungsleistung</li> <li>+ Hohes Sicherheitsniveau möglich</li> <li>+ Stabile Sensoren trotz kontaktbehaftetem Einsatz</li> </ul>
 <p>2</p>	<ul style="list-style-type: none"> <li>+ Kaum Training erforderlich</li> <li>- Ausgesprochen homogene Ausleuchtung für ICAO-Bilder nötig</li> <li>- Bei hohem Sicherheitsniveau nicht akzeptable Rückweisungsrate</li> </ul>
 <p>3</p>	<ul style="list-style-type: none"> <li>+ Hohes Sicherheitsniveau möglich</li> <li>- Bedienungsschwierigkeiten für Wenignutzer</li> <li>- FAR signifikant höher als angegeben (Mythos: feste Schwelle)</li> </ul>

**Abbildung 108: Zusammenfassender Vergleich der untersuchten Systeme**

Der Vorsprung der Fingerabdruckerkennungssysteme ergibt sich durch die insgesamt bessere Erkennungsleistung. Die Abwertung des Iriserkennungssystems resultiert aus den hohen Falschrückweisungsraten für ungeübte Nutzer.

## 10.5 *Einschätzung der Herstellerkompetenz*

Zur Einführung biometrischer Systeme für die Identitätsprüfung sind sowohl zuverlässige Systeme und Komponenten als auch kompetente Anbieter erforderlich. Ausgehend von den Erfahrungen aus dem Praxistest in BioP II werden nachfolgend einige derzeit bestehende Mankos der bereitgestellten Prototypen sowie deren Hersteller aufgezeigt.

Zum gegenwärtigen Zeitpunkt existieren nur wenige große Anwendungen biometrischer Systeme in der Praxis. Dies wird an der zum Teil mangelnden Erfahrung der Anbieter bei einem praktischen Einsatz mit einer großen Testgruppe deutlich. So existieren zumeist keine Best Practices der Hersteller hinsichtlich Einsatzumgebung und Systemkalibrierung sowie Enrolment, da im Gegensatz zu BioP viele Untersuchungen lediglich auf Basis von Datenbanken durchgeführt werden, bei denen diese Aspekte keinen Einfluss haben. Dies wird nachfolgend an einigen Beispielen deutlich:

- **Multiple Merkmalsaufnahme:** Entgegen der ursprünglichen Herstellerempfehlung müssen bei der Iriserkennung entweder beide Iriden aufgenommen werden oder es muss zumindest angezeigt werden, welche Iris zur Verifikation zu verwenden ist. Das ausschließliche Enrolment der Iris des so genannten dominanten Auges ist nicht ausreichend. Auf Basis der Untersuchungen in BioP II wird vermutet, dass bis zu 25% der Personen nicht immer das gleiche Auge zu einer Verifikation benutzen, sondern häufig wechseln.
- **Enrolment-Policy und Enrolment-Applikation:** Häufig gibt es keine klaren Empfehlungen wie ein optimales Enrolment erfolgen muss. Die Enrolment-Applikationen setzen in der Regel Fachwissen bei den Operatoren voraus. Kritischer ist jedoch, dass zum Teil keine automatische Prüfung auf eine geeignete Enrolment-Qualität und das Ausschließen von Flüchtigkeitsfehlern erfolgt. Dies wird sich insbesondere bei einem Enrolment durch nicht-technisches Personal in Meldestellen als Problem herausstellen.
- **Benutzbarkeit:** Die bereitgestellten Prototypen lassen die mangelnde Berücksichtigung und Erfahrung hinsichtlich Benutzbarkeitsaspekten erkennen. Die Benutzerführung bezüglich korrekter Bedienung und Merkmalspräsentation ist nicht ausreichend oder erfolgt verzögert. Die Bedienzeit bzw. deren subjektive Wahrnehmung wird nicht in ausreichendem Maß berücksichtigt.

Für die beispielhaft genannten Punkte besteht dringender Handlungsbedarf für die Anbieter. Diese Problemstellungen müssen bei der Auswahl eines geeigneten Anbieters für die Zielanwendung adressiert werden.

## 10.6 *Empfehlungen hinsichtlich einer Verwendung bei Personaldokumenten*

Biometrische Verfahren können die Identitätsprüfung anhand von Personaldokumenten wirksam unterstützen. Hierbei sind die unterschiedlichen Anforderungen und Umweltbedingungen verschiedener Szenarien zu beachten (z.B. verschiedene Grenzübertritte, Personenkontrolle durch die Polizei etc.). Neben technischen Maßnahmen erscheinen daneben auch organisatorische Maßnahmen geeignet, eine für die Ziele der Anwendung geeignete Erkennungsleistung in einer konkreten Anwendung unter Berücksichtigung sowohl der Falschakzeptanz- als auch der Falschzurückweisungsrate zu erreichen. Darüber hinaus sollte der Benutzer bei der Präsentation des Merkmals durch eine geeignete Benutzerführung unterstützt werden.

Für eine geeignete Bereitstellung der Referenzdaten müssen Personaldokumente um ein Speichermedium – einen kontaktlosen Chip gemäß den Vorgaben der ICAO – erweitert werden. Neben einer entsprechenden Speicherkapazität muss dieser auch über Sicherheitsmechanismen verfügen, die einen unberechtigten Zugriff auf die biometrischen Daten des Ausweisinhabers verhindern.

Eine wesentliche Voraussetzung für den erfolgreichen Einsatz biometrischer Verfahren ist die Kooperation und somit die Akzeptanz seitens der Benutzer. Aufgrund der hier gewonnenen Erkenntnisse besteht Bedarf an Aufklärung und Information sowie weiteren akzeptanzfördernden Maßnahmen.

Die Benutzerfreundlichkeit und Bedienbarkeit der hier untersuchten biometrischen Systeme ist nach den erzielten Ergebnissen verbesserungsfähig. Insbesondere müssen das Systemdesign und die Benutzerführung an der konkreten Anwendung und der Nutzergruppe ausgerichtet sein. Vor allem in Szenarien ohne zusätzliche manuelle Kontrolle sollte eine klare und einfache Benutzerführung die Bedienung des Systems so einfach und intuitiv wie möglich machen.

Beim Einsatz von Gesichtserkennung ist – wie bereits in BioP I aufgezeigt wurde – eine Anpassung der derzeitigen Aufnahme-Richtlinien zur Lichtbilderstellung erforderlich. Hier stellen die aktuellen Richtlinien der ICAO für die Erstellung von Passbildern zum Einsatz für biometrische Anwendungen eine geeignete Grundlage dar. Daneben ist unter anderem darauf zu achten, dass wesentliche Gesichtspartien nicht verdeckt sind (z. B. durch Kopfbedeckung, Frisur oder ähnliches).

Die Gesichtserkennung stellt hohe Anforderungen an die Einsatzumgebung. Hinsichtlich der Lichtbedingungen muss eine hohe Homogenität sichergestellt werden.

Weitere Verbesserungen hinsichtlich Erkennungsleistung sowie Sicherheit können möglicherweise durch geeignete Kombination verschiedener biometrischer Verfahren oder Algorithmen erreicht werden. Die Möglichkeiten und Leistungsfähigkeit solcher multimodaler Ansätze werden zurzeit geprüft.

Ebenso ist zu beachten, dass der Einfluss von Alterungseffekten auf die Erkennungsleistung biometrischer Systeme bisher noch nicht ausreichend untersucht ist. Dies ist insbesondere vor dem Hintergrund des langen Gültigkeitszeitraums von Personaldokumenten von bis zu zehn Jahren zu berücksichtigen.

Schließlich sollte zudem darauf geachtet werden, dass die Systeme hohe Anforderungen bzgl. der IT-Sicherheit sowie der physischen Sicherheit erfüllen müssen, um einen geeigneten Manipulationsschutz und Schutz der persönlichen Daten zu gewährleisten. Darüber hinaus müssen die Systeme in eine sichere Gesamtumgebung integriert werden.

Abschließend erscheint vor dem Echtbetrieb in einer konkreten Anwendung eine gründliche Untersuchung der Funktionstüchtigkeit, der Erkennungsleistung und der Überwindungssicherheit sinnvoll und notwendig. Dies kann etwa durch eine Pilotierung im Echtbetrieb unter Berücksichtigung der praktisch relevanten Parameter für die konkrete Anwendung erfolgen, um ggf. noch notwendige Anpassungen vor der tatsächlichen Inbetriebnahme in der Praxis ohne unverhältnismäßigen Aufwand vornehmen zu können.

BioP II hat das aktuelle Leistungspotenzial der drei biometrischen Verfahren Gesichtserkennung, Fingerabdruck- und Iriserkennung anhand konkreter Systeme aufgezeigt sowie Verbesserungsmöglichkeiten hinsichtlich der Erkennungsleistung einerseits und der Benutzbarkeit andererseits identifiziert. Die hier erarbeiteten Empfehlungen können als erste Richtschnur für die Verwendung biometrischer Verfahren in Verbindung mit Personaldokumenten dienen.



---

## Literatur

- [BioPI] BKA, BSI, Secunet: *BioP I - Untersuchung der Leistungsfähigkeit von Gesichtserkennungssystemen zum geplanten Einsatz in Lichtbilddokumenten*, öffentlicher Abschlussbericht 2004
- [CBEFF] ISO/IEC CD 19785-1.3 (ISO/IEC JTC 1/SC 37 N312)
- [CBEFF\_Update] ISO/IEC CD 19785-1.3 (ISO/IEC JTC 1/SC 37 N352)
- [ICAO\_BD] *BIOMETRICS DEPLOYMENT OF MACHINE READABLE TRAVEL DOCUMENTS*, ICAO TAG MRTD/NTWG, TECHNICAL REPORT; Version 1.9; 19.03.2003
- [ISO\_BPTR] ISO/IEC CD 19795-1 (ISO/IEC JTC 1 /SC 37 N684)
- [ISO\_FACE] ISO/IEC CD 19794-5 (ISO/IEC JTC 1 /SC 37 N342)
- [ISO\_FINGER] ISO/IEC CD 19794-4 (ISO/IEC JTC 1 /SC 37 N341)
- [ISO\_IRIS] ISO/IEC CD 19794-6 (ISO/IEC JTC 1 /SC 37 N339)