

DIN EN ISO 25237

ICS 35.240.80

Einsprüche bis 2015-11-18

Entwurf**Medizinische Informatik –
Pseudonymisierung (ISO/DIS 25237:2015);
Deutsche und Englische Fassung prEN ISO 25237:2015**

Health informatics –
Pseudonymisation (ISO/DIS 25237:2015);
German and English version prEN ISO 25237:2015

Anwendungswarnvermerk

Dieser Norm-Entwurf mit Erscheinungsdatum 2015-09-18 wird der Öffentlichkeit zur Prüfung und Stellungnahme vorgelegt.

Weil die beabsichtigte Norm von der vorliegenden Fassung abweichen kann, ist die Anwendung dieses Entwurfes besonders zu vereinbaren.

Stellungnahmen werden erbeten

- vorzugsweise online im Norm-Entwurfs-Portal des DIN unter www.entwuerfe.din.de bzw. für Norm-Entwürfe der DKE auch im Norm-Entwurfs-Portal der DKE unter www.entwuerfe.normenbibliothek.de, sofern dort wiedergegeben;
- oder als Datei per E-Mail an named@din.de möglichst in Form einer Tabelle. Die Vorlage dieser Tabelle kann im Internet unter www.din.de/stellungnahme oder für Stellungnahmen zu Norm-Entwürfen der DKE unter www.dke.de/stellungnahme abgerufen werden;
- oder in Papierform an den DIN-Normenausschuss Medizin (NAMED), 10772 Berlin (Hausanschrift: Burggrafenstr. 6, 10787 Berlin).

Die Empfänger dieses Norm-Entwurfs werden gebeten, mit ihren Kommentaren jegliche relevanten Patentrechte, die sie kennen, mitzuteilen und unterstützende Dokumentationen zur Verfügung zu stellen.

Gesamtumfang 145 Seiten

DIN-Normenausschuss Medizin (NAMED)

Nationales Vorwort

Dieses Dokument (prEN ISO 25237:2015) wurde vom Technischen Komitee ISO/TC 215 „Health Informatics“ (Sekretariat: ANSI, USA) in Zusammenarbeit mit dem Technischen Komitee CEN/TC 251 „Medizinische Informatik“ (Sekretariat: NEN, Niederlande) unter Beteiligung deutscher Experten erarbeitet. Von deutscher Seite ist hierfür der DIN-Normenausschuss Medizin (NAMed), Arbeitsausschuss NA 063-07-04 AA „Sicherheit“ zuständig.

Um Zweifelsfälle in der Übersetzung auszuschließen, ist die englische Originalfassung der prEN ISO 25237 beigelegt. Die Nutzungsbedingungen für den deutschen Text des Norm-Entwurfes gelten gleichermaßen auch für den englischen Text.

Für die in diesem Dokument zitierte Internationale Norm wird im Folgenden auf die entsprechende Deutsche Norm hingewiesen:

ISO 27799 siehe DIN EN ISO 27799

Nationaler Anhang NA (informativ)

Literaturhinweise

DIN EN ISO 27799, *Medizinische Informatik — Informationsmanagement im Gesundheitswesen bei Verwendung der ISO/IEC 27002*

Medizinische Informatik — Pseudonymisierung (ISO/DIS 25237:2015)

Informatique de santé — Pseudonymization (ISO/DIS 25237:2015)

Health informatics — Pseudonymisation (ISO/DIS 25237:2015)

ICS: 35.240.80

Deskriptoren:

Dokument-Typ: Europäische Norm
Dokument-Untertyp:
Dokumentstufe: parallele Umfrage
Dokumentsprache: D

STD Version 2.7g

Inhalt

	Seite
Europäisches Vorwort.....	4
Vorwort.....	5
Einleitung.....	6
1 Anwendungsbereich	7
2 Normative Verweisungen.....	7
3 Begriffe.....	7
4 Symbole (und Abkürzungen).....	14
5 Anforderungen an den Schutz von personenbezogenen Daten im Gesundheitswesen	14
5.1 ((Begriffsmodell für die Pseudonymisierung von personenbezogenen Daten))	14
5.1.1 Ziele des Schutzes von personenbezogenen Daten	14
5.1.2 Allgemeines.....	15
5.1.3 Entpersonalisierung als Prozess der Risikominderung	15
5.1.4 Schutz der personenbezogenen Daten von betroffenen Personen.....	17
5.1.5 Vergleich von personenbezogenen und entpersonalisierten Daten.....	17
5.1.6 Pseudonymisierung in der realen Welt.....	21
5.2 Kategorien der betroffenen Personen.....	24
5.2.1 Allgemeines.....	24
5.2.2 Zu betreuende Personen	24
5.2.3 Ärztliches Personal und Organisationen.....	25
5.2.4 Gerätedaten.....	25
5.3 Klassifizierung von Daten	25
5.3.1 Nutzdaten.....	25
5.3.2 Beobachtungsdaten.....	25
5.3.3 Pseudonymisierte Daten	26
5.3.4 Anonymisierte Daten.....	26
5.3.5 Forschungsdaten	26
5.3.6 Zur Identifizierung geeignete Daten	27
5.3.7 Daten von Opfern von Gewalttaten (en: Victims of Violence - VoV) und öffentlich bekannten Personen	28
5.3.8 Geninformationen.....	28
5.4 Vertrauenswürdige Dienste	29
5.5 Bedarf bezüglich der Wiedererkennung von pseudonymisierten Daten.....	29
5.6 Eigenschaften von Pseudonymisierungsdiensten.....	30
6 ((Pseudonymisierungsprozess (Verfahren und Implementierung))).....	31
6.1 Datenschutz durch Pseudonymisierung.....	31
6.1.1 Begriffsmodell der Problembereiche	31
6.1.2 Eignung von personenbezogenen Informationen zur direkten und zur indirekten Identifizierung.....	31
7 Wiedererkennungsprozess	35
Anhang A (informativ) Szenarien für die Pseudonymisierung im Gesundheitswesen	38
A.1 Einleitung.....	38
A.2 Erläuterung der Szenarien.....	38

A.3	Szenarien im Gesundheitswesen	40
A.3.1	Reihenfolge nach der klinischen Pathologie (pseudonymisierte Pflege).....	41
A.3.2	Klinische Studie.....	43
A.3.3	Klinische Forschung.....	45
A.3.4	Überwachung der öffentlichen Gesundheit.....	47
A.3.5	Berichte zur Patientensicherheit (unerwünschte Ereignisse im Zusammenhang mit Arzneimitteln)	49
A.3.6	Nicht das Gesundheitswesen betreffende Forschung, die personenbezogene medizinische Daten verwendet.....	51
A.3.7	Marktforschung.....	51
A.3.8	Lehrdateien	51
A.3.9	Außendienst.....	52
Anhang B (informativ) Anforderungen an die Risikoanalyse für den Schutz von personenbezogenen Daten		53
B.1	Einleitung	53
B.2	Bedrohungsmodell, Ziele und Mittel des Angreifers.....	55
B.3	Vollständige oder teilweise Wiedererkennung	58
B.4	Beispiel für die Wiedererkennung.....	59
B.5	Beschaffung neuer Informationen	60
B.6	Datenbankmitgliedschaft.....	61
Anhang C (informativ) Pseudonymisierungsprozess (Verfahren und Implementierung)		64
C.1	Entwicklungskriterien.....	64
C.2	Entitäten im Modell	64
C.3	Ablaufplan im Modell.....	66
C.4	Vorbereitung der Daten	67
C.5	Verarbeitungsschritte im Ablaufplan	68
Anhang D (informativ) Spezifikation von Verfahren und Implementierung		70
Anhang E (informativ) Grundstruktur einer Policy für den Betrieb von Pseudonymisierungsdiensten (Verfahren und Implementierung)		72
E.1	Allgemeines	72
E.2	Datenschutz-Policy	72
E.3	Vertrauenswürdige Praktiken für die Betriebsvorgänge.....	73
E.4	Implementierung von vertrauenswürdigen Praktiken für die Wiedererkennung	75
Literaturhinweise.....		76

Europäisches Vorwort

Dieses Dokument (prEN ISO 25237:2015) wurde vom Technischen Komitee ISO/TC 215 „Medizinische Informatik (en: Health informatics)“ in Zusammenarbeit mit dem Technischen Komitee CEN/TC 251 „Medizinische Informatik“ erarbeitet, dessen Sekretariat vom NEN gehalten wird.

Dieses Dokument ist derzeit zur parallelen Umfrage vorgelegt.

Anerkennungsnotiz

Der Text von ISO/DIS 25237:2015 wurde vom CEN als prEN ISO 25237:2015 ohne irgendeine Abänderung genehmigt.

Vorwort

ISO (die Internationale Organisation für Normung) ist eine weltweite Vereinigung von Nationalen Normungsorganisationen (ISO-Mitgliedsorganisationen). Die Erstellung von Internationalen Normen wird normalerweise von ISO Technischen Komitees durchgeführt. Jede Mitgliedsorganisation, die Interesse an einem Thema hat, für welches ein Technisches Komitee gegründet wurde, hat das Recht, in diesem Komitee vertreten zu sein. Internationale Organisationen, staatlich und nicht-staatlich, in Liaison mit ISO, nehmen ebenfalls an der Arbeit teil. ISO arbeitet eng mit der Internationalen Elektrotechnischen Kommission (IEC) bei allen elektrotechnischen Themen zusammen.

Die Verfahren, die bei der Entwicklung dieses Dokuments angewendet wurden und die für die weitere Pflege vorgesehen sind, werden in den ISO/IEC-Direktiven, Teil 1 beschrieben. Im Besonderen sollten die für die verschiedenen ISO-Dokumentenarten notwendigen Annahmekriterien beachtet werden. Dieses Dokument wurde in Übereinstimmung mit den Gestaltungsregeln der ISO/IEC-Direktiven, Teil 2 erarbeitet (siehe www.iso.org/directives).

Es wird auf die Möglichkeit hingewiesen, dass einige Elemente dieses Dokuments Patentrechte berühren können. ISO ist nicht dafür verantwortlich, einige oder alle diesbezüglichen Patentrechte zu identifizieren. Details zu allen während der Entwicklung des Dokuments identifizierten Patentrechten finden sich in der Einleitung und/oder in der ISO-Liste der empfangenen Patenterklärungen (siehe www.iso.org/patents).

Jeder in diesem Dokument verwendete Handelsname wird als Information zum Nutzen der Anwender angegeben und stellt keine Anerkennung dar.

Eine Erläuterung der Bedeutung ISO-spezifischer Benennungen und Ausdrücke, die sich auf Konformitätsbewertung beziehen, sowie Informationen über die Beachtung der WTO-Grundsätze zu technischen Handelshemmnissen (TBT, en: Technical Barriers to Trade) durch ISO enthält der folgende Link: Foreword - Supplementary information.

Das für dieses Dokument verantwortliche Komitee ist ISO/TC 215 „Health informatics“.

Einleitung

Die Pseudonymisierung gilt als wichtiges Verfahren für den Schutz von personenbezogenen Gesundheitsinformationen. Die entsprechenden Dienste können sowohl auf nationaler Ebene als auch für die grenzüberschreitende Kommunikation genutzt werden.

Zu den Anwendungsfeldern gehören u. a.:

- die indirekte Nutzung von klinischen Daten (z. B. zu Forschungszwecken);
- klinische Studien und die Überwachung nach Markteinführung;
- pseudonymisierte Pflege;
- Patientenerkennungssysteme;
- Überwachung und Bewertung der öffentlichen Gesundheit;
- vertrauliche Berichte zur Patientensicherheit (z. B. zu Nebenwirkungen von Medikamenten);
- vergleichende Berichte zu Qualitätsindikatoren;
- Kreuzgutachten (Peer-Reviews);
- Verbrauchergruppen;
- Außendienst.

Die vorliegende Technische Spezifikation stellt ein Begriffsmodell der Problembereiche sowie Anforderungen an vertrauenswürdige Praktiken und Spezifikationen zur Unterstützung der Planung und Implementierung von Pseudonymisierungsdiensten zur Verfügung.

Die Spezifikation eines allgemeinen Ablaufplans dient in Verbindung mit einer Politik bzw. Policy für vertrauenswürdige Vorgänge sowohl als allgemeiner Leitfaden für Implementierer als auch zu Qualitätssicherungszwecken, indem sie den Anwendern der Pseudonymisierungsdienste dabei hilft, zu ermitteln, inwieweit sie diesen Diensten vertrauen können. Dieser Leitfaden wird auch Ausbildungsorganisationen von Nutzen sein, indem diese befähigt werden, Pseudonymisierungsdienste selbst in einer Weise auszuführen, die ausreicht, um den gewünschten Gütegrad und die gewünschte Risikominderung zu erreichen.

1 Anwendungsbereich

Diese Technische Spezifikation enthält Grundsätze für und Anforderungen an den Datenschutz für den Fall, dass zum Schutz von personenbezogenen Gesundheitsinformationen Pseudonymisierungsdienste genutzt werden. Diese Technische Spezifikation gilt für Organisationen, die Pseudonymisierungsprozesse für ihre eigenen Zwecke ausführen wollen, und für Organisationen, die die Vertrauenswürdigkeit von an Pseudonymisierungsdiensten beteiligten Vorgängen in Anspruch nehmen.

Diese Technische Spezifikation:

- legt ein Grundkonzept für die Pseudonymisierung fest;
- gibt einen Überblick über verschiedene Anwendungsfälle der Pseudonymisierung, die sowohl umkehrbar als auch unumkehrbar sein kann;
- legt eine grundlegende Methodik für Pseudonymisierungsdienste einschließlich sowohl organisatorischer als auch technischer Aspekte fest;
- stellt einen Leitfaden für die Bewertung des Wiedererkennungsrisikos zur Verfügung;
- legt einen Rahmen für die anzuwendende Policy sowie Mindestanforderungen an vertrauenswürdige Praktiken für die Aktivitäten eines Pseudonymisierungsdienstes fest;
- legt einen Rahmen für die anzuwendende Policy sowie Mindestanforderungen an die geregelte Wiedererkennung fest.

2 Normative Verweisungen

Die folgenden Dokumente, die in diesem Dokument teilweise oder als Ganzes zitiert werden, sind für die Anwendung dieses Dokuments erforderlich. Bei datierten Verweisungen gilt nur die in Bezug genommene Ausgabe. Bei undatierten Verweisungen gilt die letzte Ausgabe des in Bezug genommenen Dokuments (einschließlich aller Änderungen).

ISO 27799, *Health informatics — Information security management in health using ISO/IEC 27002*

IHE Healthcare De-Identification Handbook: 2014

3 Begriffe

Für die Anwendung dieses Dokuments gelten die folgenden Begriffe.

3.1

Zugriffskontrolle

(en: access control)

Hilfsmittel zur Sicherstellung, dass nur entsprechend autorisierte juristische Personen auf die Ressourcen eines Datenverarbeitungssystems zugreifen können und dass dies nur auf autorisierte Art und Weise erfolgt

[QUELLE: ISO/IEC 2382-8:1998, Begriff 08.04.01]

3.2

Anonymisierung

(en: anonymization)

Prozess, bei dem personenbezogene Daten entweder vom für die Verarbeitung der Daten Verantwortlichen allein oder in Zusammenarbeit mit einer anderen Partei unumkehrbar so verändert werden, dass sich die betroffene Person danach weder direkt noch indirekt identifizieren lässt

ANMERKUNG Der Begriff ist absolut und in der Praxis möglicherweise schwierig umzusetzen.

ANMERKUNG In angepasster Form übernommen aus ISO/IEC 29100:2011 *Information technology — Security techniques — Privacy framework*.

3.3

anonymisierte Daten

(en: anonymized data)

Daten, die als Ergebnis eines Anonymisierungsprozesses erzeugt wurden

ANMERKUNG In angepasster Form übernommen aus ISO/IEC 29100:2011 *Information technology — Security techniques — Privacy framework*.

3.4

anonymer Bezeichner

(en: anonymous identifier)

Bezeichner einer Person, anhand dessen sich die betroffene natürliche Person nicht identifizieren lässt

3.5

Authentisierung

(en: authentication)

Beibringung eines Belegs für die behauptete Identität

3.6

Angreifer

(en: attacker)

Person, die versucht, mögliche Schwachstellen eines biometrischen Systems auszunutzen

[QUELLE: ISO/IEC 19792:2009 (en)]

3.7

chiffrierter Text

(en: ciphertext)

Daten, die durch Anwendung einer Verschlüsselung erzeugt wurden, deren semantischer Inhalt nicht ohne Anwendung von kryptographischen Techniken verfügbar ist

[QUELLE: ISO/IEC 2382-8:1998, Begriff 08-03-8]

3.8

Vertraulichkeit

(en: confidentiality)

Eigenschaft, dass Informationen unberechtigten natürlichen oder juristischen Personen oder Prozessen nicht verfügbar gemacht oder offenbart werden

[QUELLE: ISO 7498-2:1989, Begriff 3.3.16]

3.9

Content-Encryption-Schlüssel

(en: content-encryption key)

kryptographischer Schlüssel, der zur Verschlüsselung des Inhalts einer Kommunikation verwendet wird

3.10

für die Verarbeitung Verantwortlicher

(en: controller)

natürliche oder juristische Person, Behörde, Einrichtung oder jede andere Stelle, die allein oder gemeinsam mit anderen über die Zwecke und Mittel der Verarbeitung von personenbezogenen Daten entscheidet

3.11

Kryptographie

(en: cryptography)

Disziplin, die Grundsätze, Mittel und Verfahren für die Umwandlung von Daten verkörpert, um deren Informationsinhalte zu verbergen, deren unentdeckte Änderung und/oder deren unautorisierte Verwendung zu verhindern

[QUELLE: ISO 7498-2:1989, Begriff 3.3.20]

3.12

kryptographischer Algorithmus

Verschlüsselungsalgorithmus

(en: cryptographic algorithm)

(Chiffrierung) Verfahren für die Umwandlung von Daten mit dem Ziel, deren Informationsgehalt zu verbergen und ihre unerkannte Änderung und/oder nicht autorisierte Nutzung zu verhindern

3.13

Schlüsselverwaltung

Chiffrierschlüsselverwaltung

(en: key management, cryptographic key management)

Erzeugung, Speicherung, Verteilung, Löschung, Archivierung und Anwendung von Schlüsseln in Übereinstimmung mit einer Sicherheits-Policy (3.43)

[QUELLE: ISO 7498-2:1989, Begriff 3.3.33]

3.14

Datenintegrität

(en: data integrity)

Eigenschaft, dass Daten nicht in unautorisierter Art und Weise verändert oder zerstört wurden

[QUELLE: ISO 7498-2:1989, Definition 3.3.21]

3.15

Datenverknüpfung

(en: data linking)

Abgleich und Zusammenführung von Daten aus mehreren Datenbanken

3.16

Datenschutz

(en: data protection)

technisches und soziales Regime für die Aushandlung, Verwaltung und Sicherstellung der Vertraulichkeit und Sicherheit von Informationen

3.17

betroffene Personen

(en: data subjects)

Personen, auf die sich Daten beziehen

3.18

Dechiffrierung

Entschlüsselung

(en: decipherment, decryption)

Prozess, in dem aus einem chiffrierten Text die entsprechenden ursprünglichen Daten erhalten werden

[QUELLE: ISO/IEC 2382-8:1998, Definition 08-03-04]

ANMERKUNG Ein chiffrierter Text kann ein zweites Mal verschlüsselt werden, was dazu führt, dass eine einzelne Entschlüsselung nicht den ursprünglichen Klartext ergibt.

3.19

Entpersonalisierung

(en: de-identification)

allgemeine Benennung für jeden Prozess der Reduktion der Zuordnung zwischen einer Menge von zur Identifizierung geeigneten Daten und der betroffenen Person

3.20

zur direkten Identifizierung geeignete Daten

(en: directly identifying data)

Daten, anhand deren eine bestimmte natürliche Person direkt identifizierbar ist

ANMERKUNG Direkte Bezeichner sind diejenigen Daten, die dazu verwendet werden können, eine natürliche Person ohne zusätzliche Informationen oder durch Quervernetzung anhand anderer öffentlich zugänglicher Informationen zu identifizieren.

3.21

Offenlegung

(en: disclosure)

Preisgabe von oder Ermöglichung des Zugriffs auf Daten

ANMERKUNG Ob der Empfänger sich die Daten tatsächlich ansieht oder ob er sie zur Kenntnis nimmt oder aufbewahrt, ist ohne Bedeutung für die Frage, ob sie offengelegt wurden.

3.22

Chiffrierverschlüsselung

(en: encipherment encryption)

kryptographische Umwandlung von Daten, um **chiffrierten Text** zu erzeugen (3.6)

[QUELLE: ISO 7498-2:1989, Begriff 3.3.27]

ANMERKUNG Siehe **Kryptographie** (3.10).

3.23

Patientenbezeichner

(en: subject of care identifier, healthcare identifier)

Bezeichner einer Person für die hauptsächliche Nutzung durch ein Gesundheitssystem

3.24

identifizierbare Person

(en: identifiable person)

jemand, der direkt oder indirekt identifiziert werden kann, insbesondere über die Referenz zu einer Identifizierungsnummer oder zu einem oder mehreren Kennzeichen, die bezüglich seiner körperlichen, physiologischen, geistigen, ökonomischen, kulturellen oder sozialen Identität spezifisch sind

[QUELLE: Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr]

3.25

Identifizierung

(en: identification)

Prozess der Nutzung von behaupteten oder beobachteten Attributen einer juristischen Person mit dem Ziel, diese aus den anderen juristischen Personen in einer Reihe von Identitäten herauszufinden

ANMERKUNG Die Identifizierung einer juristischen Person in einem bestimmten Kontext ermöglicht es einer anderen juristischen Person, die juristischen Personen, mit denen sie interagiert, voneinander zu unterscheiden.

3.26

Bezeichner

(en: identifier)

Information, die vor einer möglichen Bestätigung durch einen entsprechenden Authentifikator (wie in diesem Dokument verwendet) verwendet wird, um eine Identität zu beanspruchen

[QUELLE: ENV 13608-1]

3.27

zur indirekten Identifizierung geeignete Daten

(en: indirectly identifying data)

Daten, anhand derer eine bestimmte Person nur dann identifizierbar ist, wenn sie in Verbindung mit anderen zur indirekten Identifizierung geeigneten Daten verwendet werden

ANMERKUNG Indirekte Bezeichner können die Grundgesamtheit, zu der die Person gehört, verkleinern, und zwar bei kombinierter Verwendung unter Umständen bis auf eins.

BEISPIEL Postleitzahl, Geschlecht, Alter, Geburtsdatum.

3.28

Information

(en: information)

Wissen in Bezug auf Objekte, das innerhalb eines bestimmten Kontexts eine spezielle Bedeutung hat

[QUELLE: ISO/IEC 2382-1:1993]

NEUER BEGRIFF

Daten

(en: data)

rückübertragbare, formalisierte Darstellung von Informationen, die für deren Kommunikation, Auswertung oder Weiterverarbeitung geeignet ist

ANMERKUNG Daten können von Personen oder durch automatische Hilfsmittel verarbeitet werden.

[QUELLE: ISO/IEC 2382-1:1993]

3.29

Unumkehrbarkeit

(en: irreversibility)

Situation, in der es für einen beliebigen Übergang von identifizierbar zu pseudonym rechentechnisch unmöglich ist, vom Pseudonym auf den ursprünglichen Bezeichner zu schließen

3.30

Schlüssel

(en: key)

Symbolfolge, die die Verschlüsselungs- (3.21) und Entschlüsselungsoperationen (3.17) steuert

[QUELLE: ISO 7498-2:1989, Begriff 3.3.32]

3.31

Verknüpfung von Informationsobjekten

(en: linkage of information objects)

Prozess, der die Aufstellung einer logischen Verbindung zwischen verschiedenen Informationsobjekten ermöglicht

3.32

natürliche Person

(en: natural person)

echter Mensch im Gegensatz zu einer juristischen Person, bei der es sich um eine private oder öffentliche Organisation handeln kann

3.33

Personenidentifizierung

(en: person identification)

Prozess der Aufstellung einer Verbindung zwischen einem Informationsobjekt und einer physischen Person

3.34

Personenbezeichner

(en: personal identifier)

Informationen, deren Zweck darin besteht, eine Person innerhalb eines bestimmten Kontexts eindeutig zu identifizieren

3.35

personenbezogene Daten

(en: personal data)

alle Informationen über eine bestimmte oder bestimmbare natürliche Person („betroffene Person“)

[QUELLE: Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr]

3.36

Hauptnutzung personenbezogener Daten

(en: primary use of personal data)

Nutzung von personenbezogenen Daten für die Pflege im Gesundheitswesen

3.37

Privatsphäre

(en: privacy)

Verhinderung des Eindringens in das Privatleben oder die Angelegenheiten einer Person, sofern dieses Eindringen aus der unzulässigen oder illegalen Erfassung und Nutzung von Daten über diese Person resultiert

[QUELLE: ISO/IEC 2382-8:1998, Begriff 08-01-23]

3.38

Verarbeitung personenbezogener Daten

(en: processing of personal data)

jeder mit oder ohne Hilfe automatisierter Verfahren ausgeführte Vorgang oder jede Vorgangsreihe im Zusammenhang mit personenbezogenen Daten wie das Erheben, das Speichern, die Organisation, die Aufbewahrung, die Anpassung oder Veränderung, das Auslesen, das Abfragen, die Benutzung, die Weitergabe durch Übermittlung, Verbreitung oder jede andere Form der Bereitstellung, die Kombination oder die Verknüpfung sowie das Sperren, Löschen oder Vernichten

[QUELLE: Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr]

3.39

Auftragsverarbeiter

(en: processor)

natürliche oder juristische Person, Behörde, Einrichtung oder jede andere Stelle, die personenbezogene Daten im Auftrag des für die Verarbeitung Verantwortlichen verarbeitet

[QUELLE: Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten und zum freien Datenverkehr]

3.40

Pseudonymisierung

(en: pseudonymization)

spezielle Art der Entpersonalisierung, die sowohl die ursprünglichen Identifikationsdaten zu einer Person löscht als auch die Zuordnung eines speziellen Namens (Pseudonym) zu einer Person so ermöglicht, dass anhand des zugeordneten Namens (Pseudonym) kein Rückschluss auf die Identität der Person möglich ist

3.41

Pseudonym

(en: pseudonym)

Personenbezeichner, der sich vom üblicherweise verwendeten Personenbezeichner unterscheidet und mit pseudonymisierten Daten verwendet wird, um für Kohärenz innerhalb des Datensatzes zu sorgen, die alle Informationen über die betroffene Person miteinander verknüpft, ohne die Identität dieser Person in der realen Welt offenzulegen

ANMERKUNG 1 Ein solches Pseudonym kann in umkehrbarer oder unumkehrbarer Art und Weise entweder aus dem üblicherweise verwendeten Personenbezeichner abgeleitet oder völlig ohne Zusammenhang sein.

ANMERKUNG 2 Der Begriff des Pseudonyms wird üblicherweise auf einen Bezeichner beschränkt, der eine direkte Ableitung aus dem normalen Personenbezeichner nicht zulässt. Derartige pseudonyme Informationen sind daher funktionell anonym. Unter Umständen kann eine vertrauenswürdige dritte Partei in der Lage sein, den normalen Personenbezeichner aus dem Pseudonym abzuleiten.

3.42

Empfänger

(en: recipient)

natürliche oder juristische Person, Behörde, Einrichtung oder jeder andere Stelle, die Daten erhält

3.43

sekundäre indirekte Nutzung von personenbezogenen Daten

(en: secondary indirect use of personal data)

indirekte Nutzungen sind jene, die sich vom ursprünglich vorgesehenen Verwendungszweck der erfassten Daten unterscheiden

3.44

Sicherheits-Policy

(en: security policy)

Aktionsplan oder Vorgehensweise für die Sicherstellung der Rechtersicherheit

[QUELLE: ISO/IEC 2382-8:1998, Begriff 08-01-06]

3.45

vertrauenswürdige dritte Partei

(en: trusted third party)

Sicherheitsinstanz oder deren Vertreter, die bzw. der von einer anderen Entität mit Aktivitäten in Bezug auf die Sicherheit betraut wurde

[QUELLE: ISO/IEC 18014-1:2008]

4 Symbole (und Abkürzungen)

HIPAA	Health Insurance Portability and Accountability Act
KIS	Krankenhausinformationssystem (en: Hospital Information System, HIS)
HIV	Humanes Immundefizienzvirus
IP	Internetprotokoll
VoV	Opfer von Gewalttaten (en: Victim of Violence)

5 Anforderungen an den Schutz von personenbezogenen Daten im Gesundheitswesen

5.1 ((Begriffsmodell für die Pseudonymisierung von personenbezogenen Daten))

5.1.1 Ziele des Schutzes von personenbezogenen Daten

Das Ziel des Schutzes von personenbezogenen Daten als Teil des mit der Vertraulichkeit verbundenen Ziels der Sicherheit ist, jede unbefugte oder unerwünschte Offenlegung von Informationen über eine Person zu verhindern, die in der Folge Einfluss auf rechtliche, organisatorische und/oder finanzielle Risikofaktoren haben kann. Der Schutz von personenbezogenen Daten ist ein Teilbereich des allgemeinen Datenschutzes, der der Definition nach weitere sensibel auf Verletzungen des Datenschutzes reagierende Entitäten, wie z. B. Organisationen, einschließt. Da der Schutz von personenbezogenen Daten der am besten geregelte und umfassendste ist, konzentriert sich dieses Begriffsmodell darauf. Auf den Schutz von personenbezogenen Daten ausgerichtete Lösungen lassen sich auch auf den Datenschutz anderer Entitäten bzw. betroffener Personen übertragen. Dies kann in Ländern von Nutzen sein, in denen der Schutz der Daten von betroffenen Personen oder Organisationen gesetzlich geregelt ist.

Der Schutz von personenbezogenen Daten hat zwei Ziele, nämlich einerseits den Schutz der personenbezogenen Daten beim Umgang mit Onlineanwendungen (wie z. B. Browsen im Internet) und andererseits den Schutz von in Datenbanken gespeicherten personenbezogenen Daten. Die vorliegende Technische Spezifikation befasst sich nur mit dem letzteren Ziel.

Daten können aus Datenbanken extrahiert werden. Das Ziel ist, das Risiko der Offenlegung der Identität der betroffenen Personen zu verringern. Forscher arbeiten mit „Fällen“, d. h. mit Langzeit-Patientendaten, die über bestimmte Zeiträume erfasst wurden und/oder aus verschiedenen Quellen stammen. Für die Zusammenfassung unterschiedlicher Datenelemente in derartigen Fällen ist es jedoch notwendig, ein Verfahren anzuwenden, das eine Kumulierung ohne Gefährdung des Schutzes der personenbezogenen Daten der betroffenen Personen ermöglicht. Dies lässt sich durch Pseudonymisierung der Daten erreichen.

Die Entpersonalisierung wird angewendet, um in einer Vielzahl von unterschiedlichen Situationen Risiken für den Schutz von personenbezogenen Daten zu verringern:

- Die am weitesten gehende Entpersonalisierung erfolgt bei Schulungsmaterialien, die einer breiten Öffentlichkeit zur Verfügung gestellt werden, denen sich aber dennoch Einzelheiten entnehmen lassen müssen, um für die Ausbildung von Medizinern von Nutzen zu sein. (Es gibt ein IHE-Profil für die Automatisierungsunterstützung für die Ausführung dieser Art von Entpersonalisierung. Ein großer Teil dieses Prozesses wird je nach Bedarf an den einzelnen Patienten und dem Ausbildungszweck entsprechend angepasst.)
- Das öffentliche Gesundheitswesen nutzt entpersonalisierte Datenbanken, um die Ausbreitung und den Verlauf von Krankheiten zu protokollieren und diese zu verstehen.

- In klinischen Studien wird entpersonalisiert, sowohl um personenbezogene Daten zu schützen als auch um unbewusste systematische Abweichungen zu verhindern, indem andere Informationen, wie z.B. ob der Patient ein Placebo oder ein experimentelles Medikament erhalten hat, entfernt werden.
- In geringem Maße wird in vielen klinischen Berichten entpersonalisiert, wobei dem Leser gegenüber die Identität des behandelnden Arztes, des Krankenhauses, des Patienten usw. geheim gehalten werden, um sowohl Risiken für den Schutz von personenbezogenen Daten zu verringern als auch um unbewusste systematische Abweichungen zu eliminieren. Diese Art der Entpersonalisierung verhindert nur, dass dem Leser irgendwelche Informationen zufällig offengelegt werden. Jedes planmäßige Vorgehen führt zur alsbaldigen Aufdeckung der Patientenidentität usw.
- Bei der Erstellung von Statistiken oder Analysen zur Arbeitsbelastung in Krankenhäusern oder in Bezug auf Behandlungen, die im Rahmen von Verträgen mit Beauftragten oder Kunden von Pflegediensten ausgeführt werden, ist es notwendig, einzelne Patienten aussondern zu können, ohne deren Identität im Einzelnen zu kennen. Dies ist ein Beispiel für die Entpersonalisierung in Geschäftszusammenhängen.
- Der Prozess der Risikoschichtung (z.B. erneuter Krankenhausaufenthalt) kann durch Nutzung von Datensätzen aus den primären und sekundären Patientenpflegediensten erfolgen. Die Datensätze werden für die Analyse entpersonalisiert, falls aber Patienten als hochgradig gefährdet erkannt werden, können diese von einem entsprechenden Mediziner wieder identifiziert werden, um Folgemaßnahmen einleiten zu können.

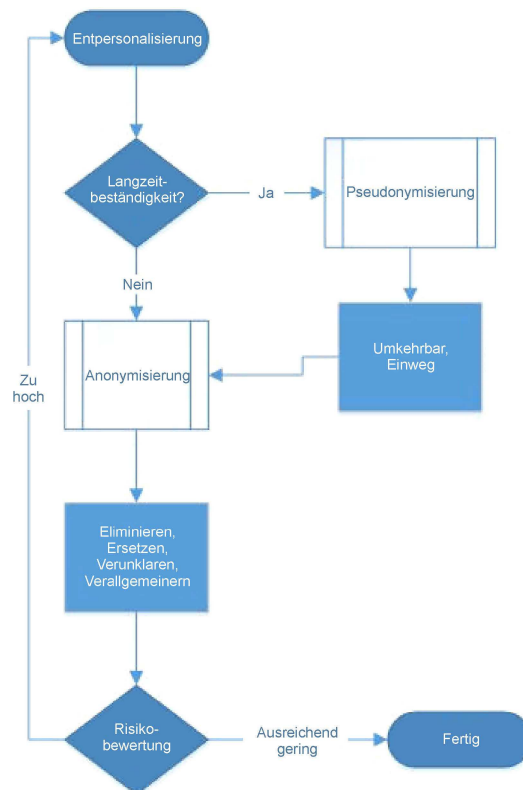
5.1.2 Allgemeines

Entpersonalisierung ist der allgemeine Begriff für jeden Prozess der Reduktion der Zuordnung zwischen einer Menge von zur Identifizierung geeigneten Daten und der betroffenen Person für einen oder mehrere vorgesehene Verwendungszwecke des resultierenden Datenbestandes. Pseudonymisierung ist eine Unterkategorie der Entpersonalisierung. Das Pseudonym ist das Mittel, mit dessen Hilfe pseudonymisierte Daten der betreffenden Person oder den betreffenden Informationssystemen verknüpft werden, ohne die Identität der Person offenzulegen. Die Pseudonymisierung kann mit oder ohne die Möglichkeit der Wiedererkennung (en: re-identification) der betroffenen Person erfolgen (umkehrbare oder unumkehrbare Pseudonymisierung). Es gibt mehrere Szenarien für die Anwendung der Pseudonymisierung im Gesundheitswesen, wobei sie besonders bei der zunehmenden elektronischen Verarbeitung von Patientendaten im Verbund mit steigenden Erwartungen der Patienten hinsichtlich des Datenschutzes Anwendung findet. In Anhang A werden mehrere Beispiele gegeben.

Es ist wichtig, zu beachten, dass über die gesamte Dauer der Existenz von pseudonymisierten Daten ein Risiko der unbefugten Wiedererkennung besteht. Insofern besteht Ähnlichkeit mit dem Verfahren der Verschlüsselung, weil sich mit genügend Aufwand letztlich jede Verschlüsselung knacken lässt, das Ziel aber ist, dies so schwierig zu gestalten, dass die damit verbundenen Kosten unerschwinglich werden. In Bezug auf die Entpersonalisierung liegen weniger Erfahrungen vor als in Bezug auf die Verschlüsselung, weswegen die damit verbundenen Risiken nicht genauso gut bekannt sind.

5.1.3 Entpersonalisierung als Prozess der Risikominderung

Im Prozess der Entpersonalisierung müssen die für den resultierenden Datenbestand geltenden Sicherheits- und Datenschutzregeln berücksichtigt werden. Nur selten lässt sich das Risiko nur durch Entpersonalisierung so weit vermindern, dass keine weiter gehenden Befugniscontrollen für den Datenbestand erforderlich sind.



Das oben stehende Diagramm ist eine bildliche Darstellung dieses Entpersonalisierungsprozesses. Es zeigt, dass an der Spitze der Begriff der Entpersonalisierung als Prozess steht. Bei diesem Prozess kommen Teilprozesse, nämlich die Pseudonymisierung und/oder die Anonymisierung, zur Anwendung. Für diese Teilprozesse wiederum werden verschiedene speziell für den betreffenden Datenelementtyp und das betreffende Risikominderungsverfahren entwickelte Tools (Hilfsprogramme) angewendet.

Der Ausgangszustand ist, dass die Menge der Daten, die das System durchlaufen dürfen, gleich null ist. Jedes Element muss anhand des vorgesehenen Verwendungszwecks des resultierenden Datenbestandes begründet werden. Der vorgesehene Verwendungszweck des Datenbestandes hat erheblichen Einfluss auf den Entpersonalisierungsprozess.

5.1.3.1 Pseudonymisierung

Falls Langzeitbeständigkeit erforderlich ist, kann die Pseudonymisierung wirksam zur Entpersonalisierung eingesetzt werden. Dies kann dazu dienen, eine Menge von Datensätzen zusammenzuhalten, die miteinander verbunden sein sollten, falls sie ohne diese Langzeitbeständigkeit voneinander getrennt werden könnten. Dies ist von Nutzen, wenn sämtliche Datensätze zu einem Patienten unter einem Pseudonym zusammengehalten werden sollen. Die Pseudonymisierung kann auch dazu dienen, sicherzustellen, dass jedes Mal, wenn Daten in einen entpersonalisierten Datensatz extrahiert werden, neue Einträge ebenfalls mit demselben Pseudonym verbunden werden. Bei der Pseudonymisierung kann der angewendete Algorithmus entweder absichtlich umkehrbar oder absichtlich unumkehrbar sein. Im Falle der Umkehrbarkeit kann eine geheime Nachschlagtabelle verwendet werden, um, sofern genehmigt, die ursprüngliche Identität aufzudecken. Im Falle der Unumkehrbarkeit kann im Verlauf des Prozesses eine temporäre Tabelle verwendet werden, die nach Abschluss des Prozesses vernichtet wird.

5.1.3.2 Anonymisierung

Bei der Anonymisierung handelt es sich um den Prozess und die Tools, die anzuwenden sind, wenn keine Langzeitbeständigkeit erforderlich ist. Der Anonymisierungsprozess wird auch dann angewendet, wenn eine Pseudonymisierung vorgenommen wurde, um die verbleibenden Datenattribute zu behandeln. Bei der Anonymisierung werden Tools, wie z.B. redaktionelle Bearbeitung, Löschen, Ausblenden, Ersetzen, Randomisieren, Verschieben, Verzerren, Kappen, Gruppieren usw., angewendet.

Für jedes Element, dem ein Durchlaufen des Systems gestattet wird, muss eine Begründung gegeben werden. Jedes Element muss das beim vorgesehenen Verwendungszweck des resultierenden Datenbestandes kleinste mögliche Risiko darstellen. Auf diese Art kann, falls der vorgesehene Verwendungszweck des resultierenden Datenbestandes keine feinkörnigen Codes erfordert, eine Gruppierung von Codes vorgenommen werden.

5.1.3.3 Direkte und indirekte Bezeichner

Im Entpersonalisierungsprozess werden drei Arten von Daten identifiziert: direkte Bezeichner, anhand derer sich der Patient direkt identifizieren lässt; indirekte Bezeichner, die bei der Anwendung in Verbindung mit anderem indirekten oder externen Wissen eine Korrelation ermöglichen; und nicht identifizierende Daten, d.h. alle anderen Daten.

Üblicherweise wird der Entpersonalisierungsprozess auf einen Datenbestand angewendet, der aus Einträgen besteht, die viele Attribute haben können; z.B. ein Arbeitsblatt mit in Spalten gegliederten Datenzeilen.

Der Entpersonalisierungsprozess wird einschließlich der Pseudonymisierung und Anonymisierung auf alle Daten angewendet. Die Pseudonymisierung wird im Allgemeinen auf direkte Bezeichner angewendet, kann aber auch auf indirekte Bezeichner angewendet werden, sofern dies für die Risikominderung bei Fortbestehen der Langzeiterfordernisse des vorgesehenen Verwendungszwecks des resultierenden Datenbestandes angemessen ist. Anonymisierungstools werden bei allen Datenformaten angewendet, sofern dies zur Risikominderung angemessen ist.

5.1.4 Schutz der personenbezogenen Daten von betroffenen Personen

Der Prozess nutzt die Vertraulichkeit von personenbezogenen Daten als Ausgangspunkt und erfordert, dass Sicherheitsmaßnahmen auf Systeme, Prozesse und Geräte angewendet werden, um die Vertraulichkeit der personenbezogenen Daten der betroffenen Person zu schützen. Die Gesetze zum Schutz der Privatsphäre enthalten Festlegungen zu einigen der in diesem Modell behandelten Begriffe. Im Kontext des Gesundheitswesens gestaltet sich der Schutz von personenbezogenen Daten wesentlich schwieriger als der Schutz der Vertraulichkeit anderer Daten, z.B. Daten zum Medikamentenverbrauch in klinischen Studien, weil sich in diesem Falle potenziell eine Vielzahl von Datentypen zur Identifizierung der betroffenen Person(en) nutzen lässt.

5.1.5 Vergleich von personenbezogenen und entpersonalisierten Daten

5.1.5.1 Definition der personenbezogenen Daten

Nach der Datenschutzrichtlinie (Richtlinie 95/46/EG des Europäischen Parlaments und des Rates vom 24. Oktober 1995 [7] (Europäische Datenschutzrichtlinie) sind „personenbezogene Daten“ *alle Informationen über eine bestimmte oder bestimmbare natürliche Person („betroffene Person“); als bestimmbar wird eine Person angesehen, die direkt oder indirekt identifiziert werden kann, insbesondere durch Zuordnung zu einer Kennnummer oder zu einem oder mehreren spezifischen Elementen, die Ausdruck ihrer physischen, physiologischen, psychischen, wirtschaftlichen, kulturellen oder sozialen Identität sind*“.

Dieser Begriff wird in anderen nationalen Gesetzen behandelt, die dieselben Grundsätze anwenden, die in dieser Definition zu finden sind (z.B. HIPAA).

5.1.5.2 Der idealisierte Begriff von Identifizierung und Entpersonalisierung

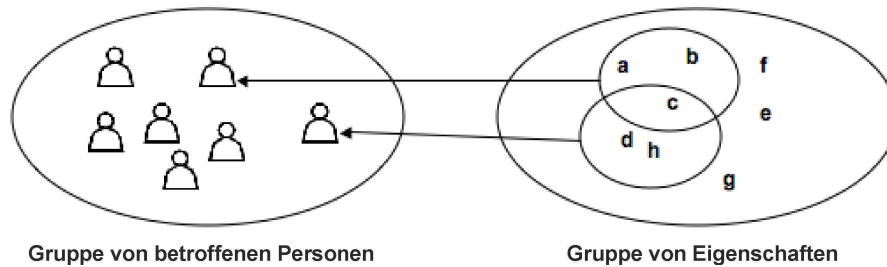


Bild 1 — Identifizierung von betroffenen Personen

In diesem Unterabschnitt wird ein idealisierter Begriff von Identifizierung und Entpersonalisierung beschrieben. Es wird davon ausgegangen, dass es außerhalb des Modells keine Daten gibt, die sich z.B. mit im Modell genutzten Daten verknüpfen lassen, um betroffene Personen (indirekt) zu identifizieren.

In 5.1.5 werden mögliche außerhalb des Datenmodells befindliche Informationsquellen berücksichtigt. Dies ist erforderlich, um die Wiedererkennungsrisiken erörtern zu können. Informations- und Kommunikationstechnologieprojekte stellen Daten, die nicht innerhalb des Modells verwendet werden, bei der Abdeckung von Aspekten des Funktionsdesigns niemals bildlich dar. Wenn sich Kritiker jedoch auf die Identifizierbarkeit konzentrieren, führen sie Informationen ein, die von einem Angreifer abgezapft werden können, um betroffene Personen zu identifizieren oder weitere Informationen über sie (z.B. zur Zugehörigkeit zu einer Gruppe) zu erhalten.

Wie in Bild 1 dargestellt, hat jede betroffene Person eine Reihe von Eigenschaften (z.B. Name, Geburtsdatum, medizinische Daten), die in medizinischen Datenbanken gespeichert werden und bei denen es sich um personenbezogene Daten der betroffenen Person handelt. Innerhalb einer Gruppe von betroffenen Personen lässt sich eine betroffene Person durch Aussondern identifizieren. Das bedeutet, dass sich eine Reihe von mit der betroffenen Person verbundenen Eigenschaften ermitteln lässt, anhand derer sich die betreffende Person eindeutig identifizieren lässt. In einigen Fällen genügt bereits eine einzige Eigenschaft, um die betroffene Person zu identifizieren (z.B. wenn es sich bei einer vorgefundenen Zahl um eine eindeutige nationale Personenkennziffer handelt). In anderen Fällen sind mehrere Eigenschaften erforderlich, um eine betroffene Person zu identifizieren, wie z.B. dann, wenn die Anschrift eines unter derselben Anschrift gemeldeten Familienmitgliedes verwendet wird. Einige Zuordnungen zwischen Eigenschaften und betroffenen Personen (z.B. Geburtsdatum, Geburtsort) sind dauerhafter als andere (wie z.B. eine E-Mail-Adresse).

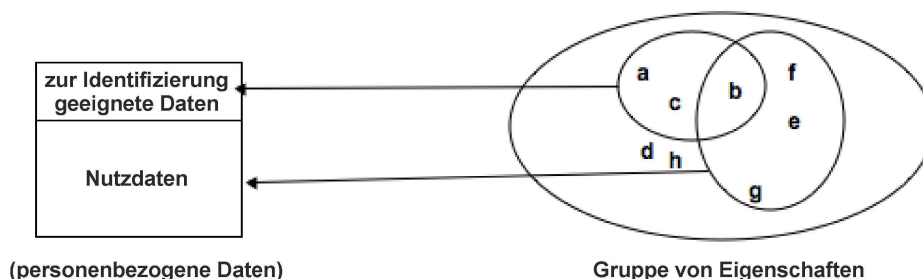


Bild 2 — Trennung der personenbezogenen Daten von Nutzdaten

Vom Begriffsstandpunkt aus lassen sich personenbezogene Daten nach den Kriterien der Identifizierbarkeit in zwei Kategorien einteilen (siehe Bild 2):

- Nutzdaten: der Teil der Daten, der Eigenschaften enthält, die keine eindeutige Identifizierung der betroffenen Person ermöglichen; begrifflich enthält die Nutzlast anonyme Daten (z.B. klinische Messwerte, maschinell erhaltene Messwerte);

- zur Identifizierung geeignete Daten: der zur Identifizierung geeignete Teil der Daten, der eine Reihe von Eigenschaften enthält, die eine eindeutige Identifizierung der betroffenen Person ermöglichen (z.B. demographische Daten).

Es ist zu beachten, dass die begriffliche Unterscheidung zwischen „zur Identifizierung geeigneter Daten“ und „Nutzdaten“ zu Widersprüchen führen kann. Dies ist dann der Fall, wenn zur direkten Identifizierung geeignete Daten als „Nutzdaten“ betrachtet werden. Jedes Pseudonymisierungsverfahren sollte darauf ausgerichtet sein, die Menge der zur direkten Identifizierung geeigneten Daten z.B. durch Zusammenfassung dieser Daten in Gruppen zu verringern. In bestimmten Fällen, in denen dies nicht möglich ist (wie z.B. beim Geburtsdatum von Säuglingen), sollte auf das damit verbundene Risiko im Policy-Dokument hingewiesen werden. In einem der folgenden Abschnitte dieses Dokuments wird die Aufteilung der Daten in den Nutzlastteil und den Teil der zur Identifizierung geeigneten Daten statt vom Begriffsstandpunkt von einem praktischen Standpunkt aus behandelt. Vom Begriffsstandpunkt aus genügt es, dass eine solche Unterscheidung möglich ist. Es ist wichtig, zu beachten, dass die Unterscheidung zwischen zur Identifizierung geeigneten Eigenschaften und Nutzlast nicht absolut ist. Einige der zur Identifizierung geeigneten Daten können auch für die Forschung notwendig sein, so z.B. das Jahr und der Monat der Geburt. Diese Unterscheidungen werden im Folgenden behandelt.

5.1.5.3 Begriff der Pseudonymisierung

Praxis und Fortschritt in der Medizin erfordern die Freigabe von Bestandteilen privater Patientenakten für die Zwecke von Lehre, Forschung und Qualitätskontrolle und für weitere Zwecke. Sowohl aus wissenschaftlichen Gründen als auch im Sinne des Schutzes von personenbezogenen Daten müssen diese Datensatzelemente modifiziert werden, um die Identitäten der betroffenen Personen zu schützen.

Es gibt kein Entpersonalisierungsverfahren, das allein in der Lage wäre, sämtliche der unterschiedlichen Erfordernisse der diversen medizinischen Nutzungen abzudecken und dabei noch die betreffenden Identitäten zu schützen. Jeder Datensatzfreigabeprozess muss einer Risikobewertung unterzogen werden, um Folgendes zu bewerten:

- a) den Zweck der Datenfreigabe (z.B. Analyse);
- b) die für diesen Zweck mindestens freizugebenden Informationen;
- c) die Risiken der Offenlegung (einschließlich Wiedererkennung);
- d) den Geheimhaltungsgrad der betreffenden Informationen (z.B. Kennzeichnung oder Etikettierung);
- e) welche Freigabestrategien zur Verfügung stehen.

Anhand dieser Informationen und der Einzelheiten des Freigabeprozesses und der Risikobewertung ist eine Strategie für die Identitätsgeheimhaltung festzulegen. Diese Festlegung muss für jeden neuen Freigabeprozess getroffen werden, obgleich für viele verschiedene Freigabeprozesse eine gemeinsame Freigabestrategie und gemeinsame Einzelheiten zur Freigabe gewählt werden können. Die meisten der in der Lehre verwendeten Dateien haben Eigenschaften den Zweck und den Mindestinformationsgehalt betreffend miteinander gemein. Vielen klinischen Arzneimittelstudien liegt bei veränderlichen Einzelheiten dieselbe Strategie zugrunde. Die Entpersonalisierung deckt mehr Erfordernisse als nur den Schutz der Vertraulichkeit ab. Häufig erfordern auch Blind- oder Doppelblindstudien eine Entpersonalisierung zur Verblindung. Dies hat Einfluss auf die Entscheidungen zu den Freigabeverfahren.

Dieser Unterabschnitt enthält die Terminologie, die zur Beschreibung der Geheimhaltung von zur Identifizierung geeigneten Informationen angewendet wird.

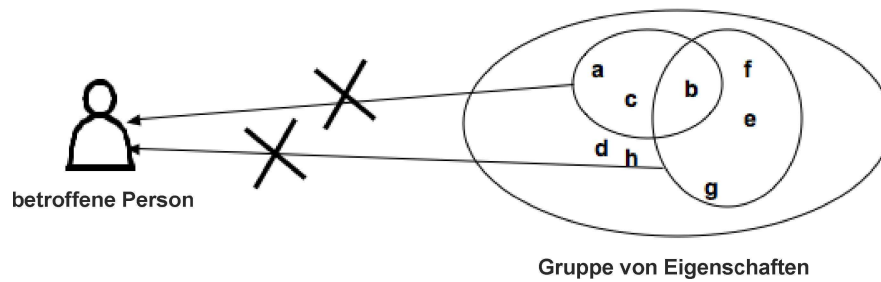


Bild 3 — Anonymisierung

Die Anonymisierung (siehe Bild 3) ist der Prozess, der die Zuordnung zwischen dem zur Identifizierung geeigneten Datenbestand und der betroffenen Person aufhebt. Dies kann auf zwei verschiedene Weisen erfolgen:

- durch Eliminierung oder Umwandlung von Eigenschaften im zugehörigen Eigenschaftsdatenbestand, so dass die Zuordnung nicht mehr eindeutig ist und sich auf mehr als eine betroffene Person beziehen kann;
- durch Vergrößerung der Grundgesamtheit der betroffenen Personen, sodass die Zuordnung zwischen dem Datenbestand und der betroffenen Person nicht mehr eindeutig ist.

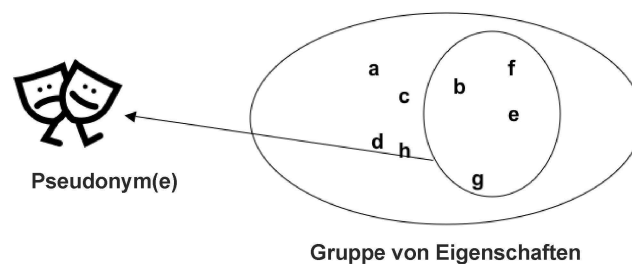


Bild 4 — Pseudonymisierung

Die Pseudonymisierung (siehe Bild 4) hebt die Zuordnung zu einer bestimmten betroffenen Person auf und fügt eine Zuordnung zwischen einem bestimmten Satz von Eigenschaften zur betroffenen Person und einem oder mehreren Pseudonymen hinzu.

Von einem funktionellen Standpunkt aus betrachtet, können pseudonyme Datenbestände zugeordnet werden, weil die Pseudonyme Zuordnungen zwischen Eigenschaftsgruppen, nicht aber die Zuordnung zur betroffenen Person ermöglichen. Auf diese Weise können z. B. Langzeitstudien durchgeführt werden, um auf realen Patientendaten basierende Fälle zusammenzutragen und die Identität der betreffenden Patienten zu schützen.

Im Falle der unumkehrbaren Pseudonymisierung enthält das Begriffsmodell kein Verfahren für die Ableitung einer Zuordnung zwischen der betroffenen Person und dem Eigenschaftsbestand aus dem Pseudonym.

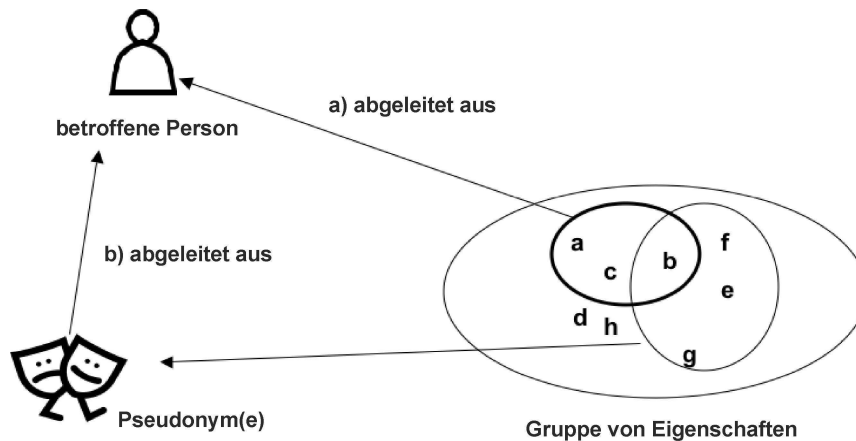


Bild 5 — Umkehrbare Pseudonymisierung

Bei der umkehrbaren Pseudonymisierung (siehe Bild 5) umfasst das Begriffsmodell eine Möglichkeit, wie sich der Datenbestand der betroffenen Person wieder zuordnen lässt.

Für dieses Ziel stehen zwei Verfahren zur Verfügung:

- Ableitung aus der Nutzlast; dies kann z.B. bei gemeinsamer Verschlüsselung von zur Identifizierung geeigneten Informationen mit der Nutzlast der Fall sein;
- Ableitung aus dem Pseudonym oder mittels einer Nachschlagetabelle.

Die umkehrbare Pseudonymisierung lässt sich auf verschiedene Weisen umsetzen, wobei davon ausgegangen wird, dass die Umkehrung der Pseudonymisierung nur unter geregelten Umständen und von einer entsprechend berechtigten Entität vorgenommen werden sollte. Das Policy-Rahmenwerk zur Wiedererkennung wird in Abschnitt 9 beschrieben. Im Gegensatz zur unumkehrbaren Pseudonymisierung erfordert die umkehrbare Pseudonymisierung üblicherweise einen erhöhten Schutz für die Entität, die die Pseudonymisierung ausführt.

Anonymisierte Daten unterscheiden sich von pseudonymisierten Daten, weil pseudonymisierte Daten ein Verfahren zur Gruppierung von Daten enthalten, das auf Kriterien beruht, die aus den personenbezogenen Daten abgeleitet werden, von denen sie abgeleitet wurden.

5.1.6 Pseudonymisierung in der realen Welt

5.1.6.1 Begründung

In 5.1.4 ist der begriffliche Ansatz für die Pseudonymisierung dargelegt, wobei Begriffe wie z.B. „zugeordnet“, „identifizierbar“, „pseudonym“ usw. als absolut angesehen werden. In der Praxis lässt sich das Risiko der Wiedererkennung von Datenbeständen häufig nur schwer beurteilen. Daher werden die Begriffe der Pseudonymisierung und der unerwünschten/unbeabsichtigten Identifizierbarkeit in diesem Abschnitt verfeinert. Als Ausgangspunkt wird an dieser Stelle die Europäische Datenschutzschutzrichtlinie in Bezug genommen.

Es gibt in vielen Rechtssystemen diverse Vorschriften, die für verschiedene Zwecke die Erstellung von entpersonalisierten Daten verlangen. Es gibt auch Vorschriften, die den Schutz von privaten Informationen verlangen, ohne die dafür anzuwendenden Mechanismen festzulegen. Diese Vorschriften verwenden in der Regel aufwands- oder schwierigkeitsbezogene Ausdrücke, was wegen des aufgrund der in rascher Entwicklung befindlichen Entpersonalisierungstechnologien veränderlichen Schwierigkeitsgrades auch angemessen ist.

Formulierungen wie „alle Mittel (...), die vernünftigerweise (...) eingesetzt werden könnten“ (en: “all the means likely reasonable”) und „von einem Dritten“ (en: “by any other person”) sind noch zu vage. Da die Definition von „identifizierbar“ und „pseudonym“ von einem nicht genau festgelegten Verhalten („alle Mittel (...), die vernünftigerweise (...) eingesetzt werden könnten“) und von nicht genau festgelegten Akteuren („von einem Dritten“) abhängt, sollte das Begriffsmodell in diesem Dokument „vernünftige“ Annahmen über „alle Mittel (...), die vernünftigerweise (...) von einem Dritten (für die Zuordnung von Eigenschaften zu betroffenen Personen) eingesetzt werden könnten“ einschließen.

Das Begriffsmodell wird künftig noch verfeinert werden, um den Unterschieden in Bezug auf die Identifizierbarkeit Rechnung zu tragen, und in das Begriffsmodell wird „Beobachtungsdatenbanken“ und „Angreifer“ einbezogen werden.

5.1.6.2 Stufen des Schutzes von personenbezogenen Daten

Den aktuellen Definitionen fehlt es bei der Beschreibung von Begriffen wie „pseudonym“ oder „identifizierbar“ an Genauigkeit. Es ist jedoch unrealistisch, anzunehmen, dass sich alle Ungenauigkeiten in der Terminologie beseitigen ließen, weil Pseudonymisierung immer auch Gegenstand von Statistik ist. Das Niveau des Risikos einer ungenehmigten Wiedererkennung lässt sich aber abschätzen. Bei der Planung der Klassifizierung dieses Risikos sollte die Wahrscheinlichkeit der Identifizierung der Fähigkeit der Daten berücksichtigt werden und ein eindeutiges Einverständnis darüber herrschen, welche betroffenen Personen in das Modell einzubeziehen sind und in welchen Beziehung zueinander diese stehen. Das Risikomodell kann in einigen Fällen auf die Minimierung des Risikos einer versehentlichen Offenlegung oder im Falle von Doppelblindstudien auf die Eliminierung einer systematischen Abweichungskomponente begrenzt werden, andererseits können die Risiken auf das Potential böswilliger Angriffe ausgedehnt werden. Das Ziel dieser Abschätzung muss sein, dass Datenschutz-Policies z.B. die „Grenzen der Ungenauigkeit“ verschieben und innerhalb eines bestimmten Kontexts festlegen können, was unter „Identifizierbarkeit“ zu verstehen ist; die Folge davon ist, dass sich Verbindlichkeiten leichter beurteilen lassen.

Im Folgenden wird eine Klassifizierung zur Verfügung gestellt, die aber noch einer weiteren Verfeinerung bedarf, besonders, weil die Quantifizierung von Wiedererkennungsrisiken die Aufstellung von mathematischen Modellen erfordert. Einen Datensatz durch einen Algorithmus laufen zu lassen, bringt unabhängig davon, wie gut der Algorithmus ist, noch immer das Risiko mit sich, wiedererkennbar zu sein. Ein kritischer Schritt im Risikobewertungsprozess ist die Analyse des resultierenden entpersonalisierten Datenbestandes in Bezug auf alle statischen Gruppen, die zur Wiedererkennung verwendet werden können. Dies ist besonders in solchen Fällen wichtig, in denen einige Bezeichner für den vorgesehenen Verwendungszweck erforderlich sind. Dieses Dokument legt keine derartigen mathematischen Modelle fest, in den Literaturhinweisen werden aber informative Quellen angegeben.

Anstelle eines idealisierten Begriffsmodells, das die außerhalb des Datenmodells befindlichen Datenquellen (ob bekannt oder unbekannt) außer Acht lässt, müssen in dem Verfahren für die Bewertung des Wiedererkennungsrisikos Annahmen getroffen werden, welche Daten außerhalb des Modells zur Verfügung stehen.

Ein praxisbezogenes Modell sollte sowohl zur direkten als auch zur indirekten Identifizierung geeignete Daten berücksichtigen. Es müssen sämtliche Anwendungsfälle analysiert werden, um die Informationsanforderungen an Bezeichner zu bestimmen und um festzulegen, welche Bezeichner einfach ausgeblendet und welche verunklart (verschwommen dargestellt) werden können und welche in voller Auflösung dargestellt und welche pseudonymisiert werden müssen.

Es werden drei Stufen des Pseudonymisierungsverfahrens festgelegt, die jeweils einen bestimmten Grad des Datenschutzes sicherstellen. Diese Sicherheitsstufen berücksichtigen die entsprechenden Wiedererkennungsrisiken basierend auf der Berücksichtigung sowohl der zur direkten als auch der zur indirekten Identifizierung geeigneten Daten. Die einzelnen Sicherheitsstufen tragen jeweils Folgendem Rechnung:

- Stufe 1: Risiken, die mit der Datenelemente identifizierenden Person verbunden sind;
- Stufe 2: Risiken, die mit der Kumulierung von Datenvariablen verbunden sind;
- Stufe 3: Risiken, die mit Ausreißern in der bestückten Datenbank verbunden sind.

Die Bewertung des Wiedererkennungsrisikos muss für alle Stufen als ständig wiederholender Prozess ausgeführt werden, wobei in regelmäßigen Abständen (wie in den Datenschutz-Policies festgelegt) erneute Bewertungen durchzuführen sind. Mit wachsender Erfahrung und immer genauerem Verständnis des Risikomodells sollten der Datenschutz und die Risikobewertungsstufen überprüft werden.

Neben den regelmäßigen erneuten Bewertungen können auch Ereignisse wie z.B. eine Änderung in den erfassten Daten oder die Einbeziehung neuer Beobachtungsdaten Anlass für eine erneute Überprüfung sein.

Bei Bezugnahme auf die Sicherheitsstufen kann die Hauptbezeichnung der Stufen 1, 2 und 3 durch die Nummer der Revision ergänzt werden (z. B. Stufe 2+ für eine einmal überarbeitete Stufe 2. Die Daten der letzten Revision sollten erwähnt werden, und es sollte ein Ereignis- und Revisionsprotokoll geführt und auf dem aktuellen Stand gehalten werden). Die gewünschte Sicherheitsstufe schreibt vor, welche Art von technischen und organisatorischen Sicherheitsmaßnahmen implementiert werden müssen, um die personenbezogenen Daten der betroffenen Person(en) zu schützen. Eine niedrige Stufe der Pseudonymisierung erfordert zum Schutz der personenbezogenen Daten eine größere Anzahl von organisatorischen Maßnahmen als eine hohe Stufe der Pseudonymisierung.

Datensicherheitsstufe 1: Eliminierung von eindeutig zur Identifizierung geeigneten Daten oder von leicht erhältlichen zur indirekten Identifizierung geeigneten Daten.

Eine erste intuitive Stufe der Anonymität lässt sich durch die Anwendung von Faustregeln erreichen. Dieses Verfahren wird bei der Erörterung von pseudonymisierten Daten üblicherweise implizit vorausgesetzt. In vielen Fällen, besonders dann, wenn nur minderbemittelte Angreifer zu erwarten sind, kann diese erste Stufe der Anonymität schon eine ausreichende Garantie sein. Zur Identifizierung geeignete Daten bedeuten, dass die in den Daten selbst enthaltenen Informationen in bestimmten Situationen bereits ausreichen, um auf eine bestimmte betroffene Person hinzuweisen. Namen von Personen sind ein typisches Beispiel. Unterabschnitt 6.6.2.2 enthält Festlegungen zu Datenelementen, die bei der Eliminierung oder Kumulierung berücksichtigt werden sollten, um einen anonymisierten Datenbestand sicherzustellen.

Datensicherheitsstufe 2: Berücksichtigung von Angreifern, die externe Daten nutzen.

Die zweite Stufe des Datenschutzes lässt sich erreichen, wenn das globale Datenmodell und die Datenströme innerhalb dieses Modells berücksichtigt werden. Bei der Festlegung der zur Erreichung dieser Stufe vorgesehenen Verfahren sollte eine statische Risikoanalyse auf Schwachstellen bezüglich der Wiedererkennung durch verschiedene Akteure durchgeführt werden. Außerdem sollte die Präsenz von möglichen Angreifern berücksichtigt werden, die externe Daten mit den pseudonymisierten Daten kombinieren, um bestimmte Datenbestände zu identifizieren. Die verfügbaren externen Daten können von der Rechtslage in verschiedenen Ländern und vom spezifischen Wissen des/der Angreifer(s) abhängen. So können die erforderlichen Verfahren z. B. die Eliminierung von absoluten Zeitangaben einschließen. Eine Bezugszeitmarke „T“ ist z. B. definiert als die Aufnahme eines Patienten zu einer bestimmten Pflegemaßnahme und andere Ereignisse; in diesem Falle wird z. B. die Entlassung in Bezug auf diese Zeitmarke ausgedrückt. Ein *Angreifer* ist eine Entität, die (mit oder ohne Genehmigung) Daten mit dem Ziel sammelt, zu versuchen, die so gesammelten Daten betroffenen Personen in nicht genehmigter Art und Weise zuzuordnen und so Informationen zu erhalten, für die sie keine entsprechende Berechtigung hat. Vom Standpunkt der Risikoanalyse werden die von einem Angreifer gesammelten und genutzten Daten als „Beobachtungsdaten“ bezeichnet.

Es ist zu beachten, dass die nicht genehmigte oder unerwünschte Aktivität des Angreifers nicht notwendigerweise das Sammeln der Daten ist, sondern vielmehr der Versuch, diese Daten einer betroffenen Person zuzuordnen und auf diese Weise Informationen zu erhalten, für die er keine entsprechende Berechtigung hat.

Ein Risikoanalysenmodell kann Annahmen über Angriffe und Angreifer einschließen. So kann es bestimmten Entitäten in einigen Ländern z. B. möglich sein, Entlassungsdaten auf legale Weise zu erhalten, sofern sie nicht implizit an der Pflege von Patienten oder der damit verbundenen Verwaltung beteiligt sind. Das Risikoanalysenmodell kann die Wahrscheinlichkeit der Verfügbarkeit von spezifischen Datenbeständen berücksichtigen.

Vom Begriffsstandpunkt aus betrachtet, bringt ein Angreifer Datenelemente in das Modell ein, die in der idealen Welt nicht existieren würden.

Ein Policy-Dokument sollte eine Beurteilung der Möglichkeit von Angriffen im gegebenen Kontext enthalten.

Datensicherheitsstufe 3: Berücksichtigung von Datenausreißern.

Das Wiedererkennungsrisiko kann in erheblichem Maße von den Daten selbst, z. B. durch das Vorhandensein von Ausreißern oder seltenen Daten, beeinflusst werden. Ausreißer oder seltene Daten können indirekt zur Identifizierung einer betroffenen Person führen. Ausreißer brauchen nicht notwendigerweise medizinische Daten zu sein. Falls z. B. an einem bestimmten Tag nur ein einziger Patient mit einer spezifischen Symptomatik eine Klinik besucht hat, können Beobachtungsdaten dazu, wer die Klinik an diesem Tag aufgesucht hat, indirekt zur Identifizierung dieses Patienten führen.

Im Falle der Beurteilung eines Pseudonymisierungsverfahrens kann eine nur statische modellbasierte Risikoanalyse aufgrund des Inhalts von Datenbanken die Anfälligkeit nicht quantifizieren; deshalb sind, um ein höheres Maß an Anonymität zu erreichen, regelmäßige Risikoanalysen an bestückten Modellen erforderlich.

In der Praxis ist die Datenschutzstufe 3 nur schwer zu erreichen.

5.2 Kategorien der betroffenen Personen

5.2.1 Allgemeines

Die vorliegende Technische Spezifikation konzentriert sich auf die Pseudonymisierung von Daten, die Patienten/zur Pflegenden betreffen. Die hierin enthaltenen Grundsätze können auch auf andere Kategorien von betroffenen Personen, wie z. B. ärztliches Personal und Organisationen, angewendet werden.

In den Unterabschnitten 5.2.2 bis 5.2.4 werden spezifische Kategorien von betroffenen Personen einzeln benannt und wird jeweils eine Reihe von mit diesen Kategorien verbundenen Problemen aufgelistet.

5.2.2 Zu betreuende Personen

Entscheidungen den Schutz der Identität der zu pflegenden Person betreffend können mit Folgendem verbunden sein:

- rechtliche Anforderungen an den Datenschutz;
- Vertrauensverhältnisse zwischen ärztlichem Personal und zu betreuenden Personen, die mit dem Grundsatz der medizinischen Schweigepflicht verbunden sind;
- verantwortungsvoller Umgang mit sensiblen Krankheitsregistern und sonstigen Informationsquellen des öffentlichen Gesundheitswesens;
- für die Erbringung von Leistungen (z. B. Labortests) mindestens erforderliche Offenlegung von Bezeichnungen;
- Datenschutz, um die indirekte Nutzung von klinischen Daten zu Forschungszwecken zu ermöglichen; dabei ist zu beachten, dass in einigen Rechtssystemen (z. B. in Deutschland) für die indirekte Nutzung von Daten zu betroffenen Personen deren Einverständniserklärung erforderlich ist, wenn die Daten nur pseudonymisiert und nicht vollständig anonymisiert sind).

Für die Kontinuität der Pflege müssen die Patienten in einheitlicher Art und Weise identifiziert werden und muss es möglich sein, Informationen über verschiedene Bereiche hinweg miteinander zu verknüpfen. Falls im Kontext der klinischen Pflege Daten pseudonymisiert werden, bestehen Risiken aufgrund von Fehlidentifizierung oder aufgrund von Fehlern bei über mehrere Bereiche erfolgende Verknüpfungen mit der betroffenen Person. Falls in einer direkten Pflegeumgebung Pseudonymisierung angewendet wird, ist die Zustimmung des Patienten für jene Fälle zu berücksichtigen, in denen der Patient aus Sicherheitsgründe keine Pseudonymisierung wünscht.

5.2.3 Ärztliches Personal und Organisationen

Pseudonymisierung kann auch angewendet werden, um die Identität von ärztlichem Personal zu schützen, dies kann u. a. folgenden Zwecken dienen:

- Kreuzgutachten;
- Bericht über medizinische Missgeschicke oder unerwünschte Ereignisse im Zusammenhang mit Arzneimitteln;
- Pflegeprozessanalyse;
- Geschäftsanalyse;
- Erstellung ärztlicher Gutachten.

Diese Schutzmaßnahmen unterliegen den rechtlichen Anforderungen der örtlichen Gesetzgebung, die sich von den Schutzanforderungen der betreffenden Organisationsidentitäten unterscheiden können.

5.2.4 Gerätedaten

Im Gesundheitswesen ist die Sicherheit der Geräte, in Unterstützung der Vertraulichkeit von Patientendaten zum Schutz der personenbezogenen Daten erforderlich. Für Patienten schließt eine Betrachtung die Berücksichtigung von implantierten Medizinprodukten ein. Zur Identifizierung geeignete Daten auf dem Gerät können sich unter Umständen dem Patienten direkt zuordnen lassen; Gleiches gilt für andere medizinische und persönliche Geräte (z.B. Beatmungsgeräte). Als solche lassen sich die Geräte-ID oder Gerätedaten zur Identifizierung einer Person nutzen. Technische Geräte im Gesundheitswesen, die einem Mitglied des ärztlichen Personals oder einem Angestellten zugewiesen sind, müssen in der Bewertung des Identifizierungsrisikos ebenfalls berücksichtigt werden, weil sie den Dienstleister oder die Organisation und somit den Patienten identifizieren können.

5.3 Klassifizierung von Daten

5.3.1 Nutzdaten

Nach dem in der vorliegenden Technischen Spezifikation befolgten Paradigma sollte es möglich sein, die Daten einzuteilen in solche, die zu einer Identifizierung führen können, und solche, die die interessierenden medizinischen Informationen tragen. Diese Bewertung ist vollständig vom angestrebten Grad des Schutzes der personenbezogenen Daten abhängig.

5.3.2 Beobachtungsdaten

Im Gesundheitswesen ist die Sicherheit der Geräte, in Unterstützung der Vertraulichkeit von Patientendaten zum Schutz der personenbezogenen Daten erforderlich. Für Patienten schließt eine Betrachtung die Berücksichtigung von implantierten Medizinprodukten ein. Zur Identifizierung geeignete Daten auf dem Gerät können sich unter Umständen dem Patienten direkt zuordnen lassen; Gleiches gilt für andere medizinische und persönliche Geräte (z.B. Beatmungsgeräte). Als solche können ein Gerät und seine Daten in der Lage

sein, eine Person zu identifizieren. Technische Geräte im Gesundheitswesen, die einem Mitglied des ärztlichen Personals oder einem Angestellten zugewiesen sind, müssen in der Bewertung des Identifizierungsrisikos ebenfalls berücksichtigt werden, weil sie den Dienstleister oder die Organisation und somit den Patienten identifizieren können. Die von einem Angreifer gesammelten und genutzten Beobachtungsdaten spiegeln verschiedene Eigenschaften von betroffenen Personen wider, die mit dem Ziel aufgezeichnet werden, die betroffenen Personen so vollständig wie möglich zu beschreiben, um ihre Zugehörigkeit zu einer bestimmten Klassifizierung in einer späteren Phase erstmalig oder erneut zu identifizieren.

5.3.3 Pseudonymisierte Daten

Es sind zwei Arten von pseudonymisierten Daten möglich.

Bei der unumkehrbaren Pseudonymisierung enthalten die pseudonymisierten Daten keine Informationen, die die Wiederherstellung der Verknüpfung zwischen den pseudonymisierten Daten und der betroffenen Person ermöglichen.

Bei der umkehrbaren Pseudonymisierung lassen sich die pseudonymisierten Daten durch Anwendung von auf entsprechend berechtigte Benutzer beschränkten Verfahren mit der betroffenen Person verknüpfen.

ANMERKUNG Die Umkehrbarkeit ist eine Eigenschaft, die sich durch verschiedene Verfahren erzielen lässt, wie z.B.: 1) gemeinsame Verschlüsselung der zur Identifizierung geeigneten Daten mit den pseudonymisierten Daten; 2) Führen einer geschützten, treuhänderisch hinterlegten Liste, die Pseudonyme mit Bezeichnern verknüpft.

5.3.4 Anonymisierte Daten

Anonymisierte Daten sind Daten, die keine Informationen enthalten, die sich zur Verknüpfung mit der betroffenen Person, zu der die Daten gehören, eignen. Eine solche Verknüpfung lässt sich z.B. durch den Namen, das Geburtsdatum, die Personenkennziffer oder andere zur Identifizierung geeignete Informationen herstellen.

5.3.5 Forschungsdaten

5.3.5.1 Allgemeines

Bei der Nutzung von Gesundheitsdaten zu Forschungszwecken handelt es sich üblicherweise um eine indirekte Nutzung von Gesundheitsdaten nach/neben ihrer Hauptnutzung zur Behandlung von Patienten. In vielen Rechtssystemen ist hierfür die Einverständniserklärung des Patienten erforderlich. Ein wesentlicher Grundsatz in Bezug auf den Schutz von personenbezogenen Daten ist, dass zur Identifizierung geeignete personenbezogene Daten nur in dem für den betreffenden Zweck erforderlichen Maße verarbeitet werden sollten. Es besteht ein eindeutiges Interesse, dass Organisationen, die Forschung betreiben, die betreffenden Daten wenn möglich pseudonymisieren oder gar anonymisieren sollten. Die Sorge um die Privatsphäre von betroffenen Personen, besonders den Bereich der gesundheitsbezogenen Informationen betreffend, hat zur Entwicklung von neuen behördlichen Anforderungen zur Sicherstellung des Schutzes von personenbezogenen Daten geführt. Forscher müssen die entsprechenden Regeln befolgen und in vielen Fällen die traditionell bei der gemeinschaftlichen Nutzung von zur Identifizierung geeigneten gesundheitsbezogenen Informationen angewendeten Verfahren entsprechend ändern.

Der Schutz von personenbezogenen Daten im Gesundheitswesen und die Selbstbestimmung des Patienten sind von entscheidender Bedeutung, viele traditionelle Schutzansätze lassen sich aber nicht ohne weiteres an die zunehmende Komplexität der Daten, der Informationsflüsse und der Möglichkeiten für zusammengeführte Informationssätze erhöhten Wertes anpassen. Es kann schwierig oder gar unmöglich sein, die klassische Einverständniserklärung für jede Art der Datennutzung zu erhalten. Im Gegensatz zu pseudonymisierten Daten jedoch kann die Forschung bei Verwendung von anonymisierten Daten fortfahren, ohne dass die betroffene Person involviert sein muss.

Trends und Gelegenheiten zur Sammlung, Zusammenführung und Wiederverwendung von zur indirekten Nutzung (z.B. zu Forschungszwecken) erfassten und gesammelten gesundheitsbezogenen Informationen werden weiter ansteigen bzw. zunehmen. Die Technologien zur Verbesserung des Schutzes von personenbezogenen Daten eignen sich gut, um den mit diesem Wachstum verbundenen Sicherheits- und Vertraulichkeitsimplikationen Rechnung zu tragen. Viele wichtige Datenanwendungen erfordern keine direkte Verarbeitung von zur Identifizierung geeigneten personenbezogenen Informationen. Wertvolle Datenanalysen lassen sich auch durchführen, ohne je die Identität der betreffenden Personen kennen zu müssen.

5.3.5.2 Generierung von Forschungsdaten

Pseudonymisierung kann auch bei der Generierung von Forschungsdaten angewendet werden. Dies ist eine optimale Gelegenheit, die mit der Forschungsstudie verbundenen Risiken für den Schutz von personenbezogenen Daten zu beurteilen und diese Risiken durch die in diesem Dokument beschriebenen Anonymisierungsverfahren zu mindern. Die Anwendungen zu Forschungszwecken erleichtern eindeutiger auch die Zustimmung und die Festlegung von Regeln, Begleitumständen und Gründen für eine absichtliche Wiedererkennung.

5.3.5.3 Indirekte Nutzung von personenbezogenen Gesundheitsinformationen

Falls nach den geltenden Gesetzen zulässig, kann Pseudonymisierung angewendet werden, um die personenbezogenen Daten von Personen zu schützen, deren personenbezogene Gesundheitsinformationen zu indirekten Zwecken genutzt werden sollen. Indirekte Nutzungen sind solche, die sich vom ursprünglich vorgesehenen Verwendungszweck der erfassten Daten unterscheiden. Für jede indirekte Nutzung müssen die damit verbundene Bedrohung für den Schutz der personenbezogenen Daten bewertet und Maßnahmen zur Minderung der ermittelten Risiken festgelegt werden. Es dürfen keine Annahmen bezüglich der Zulänglichkeit einer bereits erstellten Risikobewertung und ergriffener Risikominderungsmaßnahmen getroffen werden, um die Datenressource auf weitere indirekte Nutzungen auszudehnen.

5.3.6 Zur Identifizierung geeignete Daten

Daten, die Informationen enthalten, die eine eindeutige Identifizierung der betroffenen Person ermöglichen (z. B. demographische Daten).

5.3.6.1 Patientenbezeichner

Im Gesundheitswesen sollten widersprüchliche identitätsbezogene Anforderungen einander angeglichen werden.

- Sofern eine entsprechende Genehmigung vorliegt, dürfen mehrere medizinische Datenquellen, die sich auf eine benannte betroffene Person beziehen, über verschiedene Bereiche hinweg miteinander verknüpft werden. Je nach den Nutzungsanforderungen für die miteinander verknüpften Daten kann eine
 - korrekte Verknüpfung (keine Verknüpfung von Datenquellen, die sich auf unterschiedliche Patienten beziehen);
 - vollständige Verknüpfung (keine fehlenden Verknüpfungsglieder, weil dies zu Fehlern bei der Identifizierung der betroffenen Person führen kann)
- erforderlich sein.
- Falls der Zugriff auf die zur Identifizierung der betroffenen Person geeigneten Daten beschränkt ist, können die Daten von entsprechend berechtigten Stellen unter geregelten Umständen und mit Hilfe eines vertrauenswürdigen Dienstleisters (Service Provider) mit den betroffenen Personen verknüpft werden.

In einigen Rechtssystemen sind Verknüpfungen zwischen verschiedenen Bereichen nur begrenzt zulässig. Dem muss ebenfalls Rechnung getragen werden. Falls eine betroffene Person verschiedene Anbieter von gesundheitsbezogenen Dienstleistungen aufgesucht hat, verwenden diese Anbieter häufig nur ihre eigene interne Benummerung. Verwaltungstechnische und medizinische Informationen werden anderen Stellen häufig mit diesen lokal vergebenen Nummern übergeben. Dies hat zur Folge, dass Stellen, die kumulierte Daten benötigen, nicht sicher sein können, dass die kumulierten Daten auch vollständig sind.

Dies lässt sich vermeiden, indem bei der Identitätsverwaltung ein strukturierter Ansatz angewendet wird. Es stehen mehrere Ansätze für die Identitätsverwaltung zur Verfügung, weswegen eine ausführliche Erörterung der Identitätsverwaltung nicht Gegenstand der vorliegenden Technischen Spezifikation ist. Dennoch hat eine Reihe von Identitätsverwaltungslösungen eine Pseudonymisierungslösung zum Kern.

5.3.7 Daten von Opfern von Gewalttaten (en: Victims of Violence - VoV) und öffentlich bekannten Personen

Bei der Untersuchung und Behandlung von Opfern von Gewalttaten ist häufig eine zusätzliche Abschirmung durch das Krankenhauspersonal erforderlich; dies gilt, solange mit ihrer Identifizierung eine spezielle Bedrohung verbunden ist. Pflegepersonal, das in direktem Kontakt mit dem Patienten steht, kann die Person identifizieren, das Verwaltungspersonal dagegen kann das nicht.

Ähnliche Probleme treten häufig auf, wenn öffentlich bekannte Personen oder dem ärztlichen und Pflegepersonal auf andere Weise bekannte Personen (häufig fälschlich als „VIPs“ [Very Important Persons] oder Prominente bezeichnet) aufgenommen werden (z. B. Politiker, führende Industrielle usw.).

5.3.8 Geninformationen

Hinsichtlich Geninformationen gibt es zwei Denkschulen. Der eine Standpunkt lautet, dass Geninformationen etwas anderes als alle anderen medizinischen Informationen seien. Der andere Standpunkt lautet, dass die Gendaten in gleicher Art und Weise zu betrachten sind wie jede andere medizinische Information. Für den Standpunkt, dass sich Gendaten von allen anderen medizinischen Daten unterscheiden, wurde der Begriff „genetischer Exzeptionalismus“ geprägt.

2004 hat die Europäische Kommission Direktion C (Wissenschaft und Gesellschaft), Referat C3 (Ethik und Wissenschaft) 25 Empfehlungen zu den ethischen, rechtlichen und sozialen Fragen von Gentests herausgegeben. Diese Empfehlungen lauten wie folgt:

- „genetischer Exzeptionalismus‘ sollte vermieden werden, sowohl auf internationaler Ebene als auch auf Ebene der EU und ihrer Mitgliedstaaten. Zugleich ist es notwendig, die öffentliche Wahrnehmung, wonach Gentests ‚etwas anderes‘ seien, anzuerkennen und zu diskutieren;
- alle medizinischen Daten, einschließlich Gendaten, sollten gleich hohen Qualitäts- und Vertraulichkeitsstandards genügen müssen.“

Bezüglich der Vertraulichkeit, Privatheit und Selbstbestimmung lauten die Empfehlungen:

- „Gendaten von klinischer und/oder familiärer Bedeutung sollten in gleichem Maße geschützt werden wie andere gleichermaßen sensible medizinische Daten;
- die Relevanz der Daten für andere Familienmitglieder sollte berücksichtigt werden;
- die Bedeutung des Rechts des Patienten auf Wissen oder Nichtwissen so geachtet und praktische Mechanismen eingeführt werden, die gewährleisten, dass dieses Recht respektiert wird;
- Im Kontext von Gentests sollten Verfahrensweisen erarbeitet werden hinsichtlich der Bereitstellung von Informationen, der Beratung, der Einholung einer informierten Einwilligung und der Bekanntgabe der Testergebnisse, um diesen Anforderung auch gerecht zu werden.“

Es gibt natürlich auch andere Meinungen, die nicht mit dieser Sichtweise übereinstimmen. Ein Argument ist z. B., dass der Informationsgehalt von im Rahmen der Bevölkerungsforschung erfassten Geninformationen nicht genau bekannt ist und dass der Informationsgehalt von in Datenbanken oder menschlichen Gewebesammlungen gespeicherten Daten sensibler sein kann, als sich gegenwärtig abschätzen lässt.

Andere Gruppen folgen diesem Gedankengang nicht, so z. B. die Unterzeichner der als „Erklärung von Montreux“ bekannt gewordenen Schlussklärung der im September 2005 abgehaltenen 27. Internationalen Konferenz der Beauftragten für den Datenschutz und den Schutz der Privatsphäre“, in deren Präambel es heißt: (Die Datenschutzbeauftragten sind) „sich bewusst, dass aufgrund des rasch wachsenden Kenntnisstandes im Bereich der Genetik Daten über die menschliche DNA zu den sensibelsten überhaupt werden können und dass die Gewährleistung eines angemessenen rechtlichen Schutzes dieser Daten angesichts der beschleunigten Wissensentwicklung wachsende Bedeutung erlangt“.

Dennoch müssen, wann immer Geninformationen Bestandteil eines zu pseudonymisierenden Datenbestandes sind, die bestehenden Gesetze, Richtlinien und Veröffentlichungen zu den damit verbundenen Themen berücksichtigt werden, die sich häufig mit dem weiter gefassten Rahmen der auf ganze Populationen bezogenen Genforschung und mit Sammlungen von menschlichem Gewebe befassen. Allgemein ausgedrückt, dürfen Datenressourcen entweder als Identifizierung von Krankheitsgenen oder von diagnostischen Biomarkern klassifiziert werden.

5.4 Vertrauenswürdige Dienste

Für den Fall, dass eine Synchronisierung der Pseudonyme durch den Pseudonymisierungsdienst über mehrere Entitäten oder Unternehmen hinweg erforderlich ist, darf ein vertrauenswürdiger Dienstanbieter beauftragt werden. Für die Implementierung vertrauenswürdiger Dienste stehen diverse Optionen zur Verfügung, einschließlich der Beauftragung kommerzieller Einrichtungen, Mitgliedsorganisationen oder staatlicher Stellen. Vertrauenswürdige Dienstleister können in den verschiedenen Rechtssystemen gesetzlichen oder Zertifizierungsanforderungen unterliegen.

5.5 Bedarf bezüglich der Wiedererkennung von pseudonymisierten Daten

Bei der Pseudonymisierung werden die zur Identifizierung von Personen geeigneten Daten aus den Nutzdaten ausgesondert, indem den sensiblen Daten vor dem Aussplittern der Daten ein codierter Wert zugewiesen wird. Beim umkehrbaren Ansatz wird zwischen den Nutzdaten und den Personenbezeichnern eine Verbindung aufrechterhalten, dies kann aber unter bestimmten vorgeschriebenen Umständen und bei Einhaltung entsprechender Schutzmaßnahmen eine Wiedererkennung ermöglichen. Beim unumkehrbaren Ansatz wird keine Verbindung zwischen Nutzdaten und Personenbezeichnern beibehalten, und entsprechend ist keine Wiedererkennung möglich.

Dieser Ansatz leistet den Forschern gute Dienste, weil er eine Möglichkeit zur Bereinigung der Forschungsdaten bietet, während die Fähigkeit zur Referenzierung von Quellenbezeichnern für die vielen (geregelten) Umstände erhalten bleibt, unter denen solche Informationen erforderlich sein können. Zu diesen Umständen gehören die folgenden codierten Werte; Vokabular-Identifizierung: ISO (1) Norm (0) Pseudonymisierung (25237) Zweck der Wiedererkennung (1):

- 1) Verifizierung/Validierung der Datenintegrität;
- 2) Verifizierung/Validierung des Datensatzdoppels;
- 3) Bitte um weitere Daten;
- 4) Verknüpfung mit ergänzenden Informationsvariablen;
- 5) Audit zur Prüfung auf Einhaltung der Anforderungen;
- 6) Mitteilung bedeutsamer Befunde;
- 7) Folgeuntersuchungen.

Diese Werte sollten für Auditzwecke wirksam eingesetzt werden, wenn es darum geht, die berechnete Wiedererkennung zu erleichtern. Die angewendeten Wiedererkennungsverfahren müssen gründlich abgesichert werden und können in Form der Nutzung eines vertrauenswürdigen Dienstes für die Generierung und Verwaltung der Dechiffrierschlüssel umgesetzt werden. Die Kriterien für die Wiedererkennung können mit Hilfe der vertrauenswürdigen Dienste festgelegt, automatisiert und sicher verwaltet werden.

5.6 Eigenschaften von Pseudonymisierungsdiensten

Es gibt zwei Hauptszenarien für Pseudonymisierungsdienste:

- 1) Die Pseudonyme werden innerhalb einer einzelnen Organisation oder für diese oder zu genau einem bestimmten Zweck verwaltet – in diesem Fall befasst sich der Dienst in der Regel mit Identitäten, die der Organisation zugewiesen wurden oder ihr bekannt sind.
- 2) Die Pseudonyme werden von dem oder den Pseudonymisierungsdiensten bereitgestellt – in diesem Fall stellt der Dienst in der Regel Pseudoidentitäten zur Verfügung, die für nicht angeschlossene Organisationen gelten und die Verknüpfung von gesundheitsbezogenen Patienteninformationen ermöglichen und dabei die Identität dieser Patienten schützen.

In beiden Fällen muss der Dienst so erbracht werden, dass das Risiko einer nicht genehmigten Wiedererkennung der vom Pseudonymisierungsdienst betroffenen Personen auf ein Mindestmaß reduziert wird.

Der mit dem Schutz der Patientenidentitäten betraute Dienst muss in seinen Praktiken die Mindestanforderungen an die Vertrauenswürdigkeit erfüllen:

- Es ist notwendig, das Vertrauen der Patienten in die Fähigkeit des Gesundheitssystems zur Wahrung der Vertraulichkeit ihrer Informationen zu stärken;
- der Dienst muss die physische Sicherheit gewährleisten;
- der Dienst muss die Betriebssicherheit gewährleisten;
- Re-Identifizierungsschlüssel, Umsetzungstabellen und der Schutzbedarf müssen Kontrollen durch jeweils mehrere Personen und/oder mehrere Organisationen unterzogen werden, die den vom Dienst gegebenen Zusicherungen entsprechen;
- der Dienst muss unter der (z. B. vertraglichen oder betrieblichen) Kontrolle der Stelle stehen, bei der die Quellenbezeichner hinterlegt sind;
- rechtliche und umweltbezogene Einschränkungen bezüglich der Freigabe von Re-Identifizierungsschlüsseln und Schutzmaßnahmen müssen zur Unterstützung der vom Dienst angebotenen Datensicherheitsstufen offengelegt werden;
- die Dienstgüte und -verfügbarkeit müssen festgelegt und dem Bedarf bezüglich der Informationsbereitstellung und des Informationszugriffs entsprechend geboten werden;
- einige Bezeichner können einfach ausgeblendet werden, weil sie für die Nutzung nicht benötigt werden;
- einige Bezeichner dürfen auch verschwommen dargestellt werden, sofern dies dem vorgesehenen Verwendungszweck entspricht.

6 ((Pseudonymisierungsprozess (Verfahren und Implementierung)))

6.1 Datenschutz durch Pseudonymisierung

6.1.1 Begriffsmodell der Problembereiche

Die vorliegende Technische Spezifikation gilt in erster Linie für gesammelte oder gespeicherte Informationen, weniger für die interaktive Nutzung von Systemen durch Patienten. (Vom Patienten im Verlauf einer solchen interaktiven Nutzung eingegebene oder bearbeitete Informationen können als gespeicherte Information angesehen werden.)

Es gibt mehrere Gründe dafür, personenbezogene Daten durch Geheimhaltung der betreffenden Identitäten zu schützen. In jedem Fall müssen in der betreffenden Datenschutz-Policy Ziele für den Schutz von personenbezogenen Daten durch Pseudonymisierung festgelegt werden, wobei anzugeben ist, was als zur Identifizierung geeignete Information angesehen wird und was nicht.

Vom Funktionalitätsstandpunkt aus betrachtet, ist es wichtig, anzugeben, ob Umkehrbarkeit erforderlich ist und was die Endergebnisse der Umkehrbarkeit sind, um die berechtigte Anwendung der Umkehrbarkeit verfahrensmäßig und technisch zu erleichtern, während andere Anwendungen verhindert werden.

Im Rahmen der Identitätsverwaltung können komplexe Pseudonymisierungsfunktionen erforderlich sein, zu denen je nach Identitätsverwaltungsschema die Übertragung von Pseudonymen zwischen Identitätsbereichen gehören kann.

Zwei wichtige Elemente des Konzepts der Pseudonymisierung sind:

- der Bereich, in dem ein Pseudonym verwendet wird;
- Schutz des Pseudonymisierungsschlüssels oder des „Samenkorns“ (en: seed).

6.1.2 Eignung von personenbezogenen Informationen zur direkten und zur indirekten Identifizierung

6.1.2.1 Allgemeines

Personenbezogene Daten können sich zur direkten oder der indirekten Identifizierung eignen. Die Daten gelten als zur direkten Identifizierung geeignet, wenn es möglich ist, anhand eines Datenattributs oder durch Verknüpfung dieses Attributs mit einer öffentlich zugänglichen Ressource oder mit einer zugriffsbeschränkten Ressource unter einer alternativen Policy, die die Identität enthält, eine bestimmte Person zu identifizieren. Hierzu würden Querverweise auf bekannte Bezeichner (wie z.B. Telefonnummer, Anschrift) oder numerische Bezeichner (z.B. Auftragsnummern, Studiennummern, Dokument-Objektbezeichner, Zahlenwerte von Laborergebnissen) gehören. Ein indirekter Bezeichner ist ein Attribut, das in Verbindung mit zur indirekten Identifizierung geeigneten Attributen verwendet werden kann, um die betreffende Person eindeutig zu identifizieren (z.B. Postleitzahl, Geschlecht, Geburtsdatum). Dies schließt auch geschützte zur indirekten Identifizierung geeignete Bezeichner (z.B. Behandlungsdatum, Datum der Bilderstellung) ein, für die die Zugriffsmöglichkeiten stärker eingeschränkt sind, die aber zur Identifizierung des Patienten genutzt werden können.

6.1.2.2 Zur Identifizierung von Personen geeignete Variablen

Zu den zur Identifizierung von Personen geeigneten Variablen gehören:

- Personennamen (einschließlich Rufname, bürgerlicher Name, sonstige Namen, unter denen die Person bekannt ist); mit dem Namen beziehen wir uns auf den Namen und allen Namensdatenelemente, wie in ISO/TS 22220 festgelegt;
- Personenbezeichner (einschließlich z.B. ausstellende Behörden, Typen und Bezeichnungen wie Patientenabrechnungsnummer, Nummer der Patientenakte, Zertifikats-/Zulassungsnummern, Sozialversicherungsnummer, Nummern der Empfangsberechtigten von Krankenkassenleistungen des Arbeitgebers, Fahrzeugbezeichner und Seriennummern, einschließlich amtlicher Kfz-Kennzeichen);
- biometrische Kennzeichen (Stimmprofile, Fingerabdrücke, Fotos usw.);
- digitale Zertifikate, die sich zur Identifizierung einer Person eignen;
- Geburtsname der Mutter und sonstige ähnliche beziehungsbasierte Konzepte (z.B. Familienbeziehungen);
- Wohnanschrift;
- elektronische Kommunikationswege (Telefon, Mobiltelefon, Telefax, Pager, E-Mail, URL, IP-Adressen, Gerätebezeichner und Geräteseriennummern);
- Verbindungen der betroffenen Person (Mutter, Vater, Geschwister, Kind);
- Beschreibungen von Tätowierungen und persönlichen Erkennungszeichen.

Je nach angewendetem Datenformatstandard können entsprechende Standardspezifikationen zur Verfügung stehen, die befolgt werden sollten (z.B. DICOM PS 3.15 Annex E).

6.1.2.3 Kumulationsvariablen

Für statistische Zwecke sollten absolute Datenverweisungen vermieden werden.

- Geburtsdaten eignen sich z.B. besonders gut zur Identifizierung. Altersangaben, eignen sich weniger gut, können aber immer noch eine Bedrohung durch Verknüpfung von Beobachtungsdaten darstellen, weswegen es besser ist, Altersgruppen oder Alterskategorien zu verwenden. Um sichere Bereiche zu bestimmen, sollte eine Analyse bezüglich des Wiedererkennungsrisikos durchgeführt werden, die aber nicht Gegenstand der vorliegenden Technischen Spezifikation ist.
- Aufnahme- und Entlassungsdaten usw. lassen sich ebenfalls in Kategorien von bestimmten Zeiträumen zusammenfassen, Ereignisse könnten aber jeweils in Bezug auf einen bestimmten Meilenstein ausgedrückt werden (z.B. x Monate nach der Behandlung).
- Ortsdaten sollten, falls regionsbezogene Codes zu spezifisch sind, kumuliert werden. Falls Ortscodes hierarchisch gegliedert werden, können die feineren Abstufungen weggelassen werden, z.B. falls Postleitzahlen oder Vorwahlnummern für 20 000 oder weniger Menschen gelten, darf der entsprechende Code in 000 geändert werden.¹⁾

¹⁾ HIPAA Abschnitt 164.514

Bei demographischen Daten kann es sich sowohl um direkte als auch um indirekte Bezeichner handeln, und sie sollten wenn möglich entfernt werden; ist dies nicht möglich, so sollten sie ab einem durch den Bereich oder das Rechtssystem festgelegten Schwellenwert kumuliert werden. Falls diese Daten aufbewahrt werden müssen, muss eine Bewertung des Risikos einer nicht genehmigten Wiedererkennung vorgenommen werden und sind die entsprechenden Maßnahmen zur Minderung dieses Risikos für die resultierende Datenressource durchzuführen. Zu derartigen demographischen Daten gehören:

- die zu Hause gesprochene Sprache;
- die Kommunikationssprache der Person;
- die Religion;
- die Volkszugehörigkeit;
- das Geschlecht der Person;
- das Geburtsland;
- der Beruf;
- Vorstrafen;
- Rechtsordnungen der Person;
- andere Anschriften (z. B. Geschäftsanschrift, zeitweise Anschriften, Postanschriften);
- Mehrling (Nachgeborener in einer Zwillings- oder Mehrlingsgeburt).

Es muss ein Policy-Dokument erstellt werden, das eine Bewertung der Möglichkeit von Angriffen im gegebenen Kontext als Risikobewertung in Bezug auf die Datenschutzstufe 2 enthält. Für die ermittelten Risiken muss eine entsprechende Risikominderungsstrategie angegeben sein.

6.1.2.4 Ausreißervariablen

Basierend auf einer Risikobewertung sollten Ausreißervariablen entfernt werden.

Ausreißervariablen enthalten:

- seltene Diagnosen;
- ungewöhnliche Behandlungen;
- einige Berufe (z. B. berufsmäßiger Tennisspieler);
- bestimmte rezessive Merkmale, die für die betreffende Population der Informationsquelle untypisch sind;
- deutliche Fehlbildungen.

Es muss ein Policy-Dokument erstellt werden, das eine Bewertung der Möglichkeit von Angriffen im gegebenen Kontext als Risikobewertung in Bezug auf die Datenschutzstufe 3 enthält. Für die ermittelten Risiken muss eine entsprechende Risikominderungsstrategie angegeben sein.

Persistierende Datenressourcen, die Pseudonymität beanspruchen, müssen einer routinemäßigen Risikoanalyse auf potentiell zur Identifizierung geeignete Ausreißervariablen unterzogen werden. Diese Risikoanalyse muss mindestens einmal jährlich durchgeführt werden. Für die ermittelten Risiken muss eine entsprechende Risikominderungsstrategie angegeben sein.

6.1.2.5 Strukturierte Datenvariablen

Strukturierte Daten geben einen Hinweis darauf, welche Informationen erwartet werden können und wo diese zu erwarten sind. Danach obliegt es der Analyse des Wiedererkennungsrisikos, Annahmen darüber zu treffen, was zu (unannehmbaren) Identifizierungsrisiken führen kann; das reicht von einfachen Faustregeln bis zur Analyse von bestückten Datenbanken und Inferenz-Deduktionen. Bei „freiem Text“ ist im Gegensatz zu „strukturiert“ keine automatisierte Analyse für den Zweck des Schutzes von personenbezogenen Daten mit garantiertem Ergebnis möglich.

6.1.2.6 Nicht-strukturierte Datenvariablen

Im Falle von nicht-strukturierten Datenvariablen bleibt die Pseudonymisierungsentscheidung der Einteilung der Daten in zur Identifizierung geeignete Daten und Nutzdaten die Kernfrage. Freiform-Text ist als verdächtig anzusehen, und es sollte daher erwogen werden, ihn zu eliminieren. Nicht-strukturierte Datenvariablen müssen Folgendem unterzogen werden:

- Aussondern, was nach der Datenschutz-Policy (und der gewünschten Datenschutzzstufe entsprechend) zur Identifizierung geeignete Informationen sind;
- Löschen von Daten, die nicht benötigt werden;
- die Policies sollten angeben, dass der Freitextteil keine zur direkten Identifizierung geeigneten Informationen enthalten darf.

Die Teile, die nach der Policy als nicht zur Identifizierung geeignet angesehen werden, werden als Nutzdaten zusammengehalten.

Freiform-Text

Bei aktuellen Pseudonymisierungsansätzen kann für Freiform-Text keine Anonymität sichergestellt werden. Freiform-Text muss in jedem Falle einer Risikoanalyse unterzogen werden, und für die ermittelten Risiken muss eine entsprechende Risikominderungsstrategie angegeben sein. Die Wiedererkennungsrisiken des beibehaltenen Freiform-Texts lassen sich durch Folgendes mindern:

- Implementierung einer Policy betreffend den Inhalt des Freiform-Textes, die erfordert, dass die Freiform-Textdaten keine zur direkten Identifizierung geeigneten Informationen (z.B. Patientenkennummern, Namen) enthalten dürfen;
- Verifizierung, dass der Freiformgehalt wahrscheinlich keine zur Identifizierung geeigneten Daten enthält (z.B. falls Freiformtext aus strukturiertem Text erzeugt wird);
- Überarbeitung, Umschreiben oder sonstige Umwandlung der Daten in codierte Form.

Mit der fortschreitenden Entwicklung von Parsing(Syntaxanalyse-)algorithmen und von Datenbereinigungs- und Pseudonymisierungsalgorithmen, die sich zur Verarbeitung von natürlicher Sprache eignen, kann diese Aussage aufgrund der mit Freiformtext verbundenen Wiedererkennungsrisiken möglicherweise gelockert werden.

Freiformtext sollte überarbeitet, umgeschrieben oder auf andere Weise in Code umgewandelt werden.

Text-/Stimmdaten mit nicht-syntaktisch analysierbarem Inhalt

Wie bei Freiformtext sollten nicht-syntaktisch analysierbare Daten, wie z.B. Stimmfelder, eliminiert werden.

Bilddaten

Einige medizinische Daten enthalten zur Identifizierung geeignete Informationen (z.B. Röntgenbilder, auf denen sich Patientenbezeichner erkennen lassen). Die Abhilfemaßnahmen für derartige zur Identifizierung geeignete Daten in der strukturierten und codierten DICOM-Kopfzeile sollten DICOM PS3.15 Annex E entsprechen. In DICOM (ISO 12052) sind empfohlene Entpersonalisierungsprozesse für DICOM SOP-Instanzen (Dokumente) für einige häufig anzutreffende Situationen festgelegt. Außerdem ist darin eine Liste von verschiedenen anwendbaren Entpersonalisierungsalgorithmen festgelegt. Darüber hinaus werden darin einige häufig anzutreffende Anwendungssituationen und -eigenschaften benannt, z.B. "need to retain device information" (etwa: die Notwendigkeit, die Geräteerkennung beizubehalten). Für jedes Standard-DICOM-Attribut (Datenelement) wird der Algorithmus empfohlen, der sich für das betreffende Attribut in der jeweiligen Situation mit hoher Wahrscheinlichkeit am besten eignet.

Es wird erwartet, dass diese Zuweisungen, sobald dies angemessen ist, angepasst werden, die Angabe eines Ausgangspunktes für typische Situationen verringert den mit der Festlegung eines Entpersonalisierungsprozesses verbundenen Arbeitsaufwand aber schon erheblich. Für zur Identifizierung geeignete Eigenschaften des Bildes oder der Beschriftungen, die Teil des Bildes sind, muss eine zusätzliche Risikobewertung erwogen werden.

6.1.2.7 Inferenzrisikobewertung

Es ist anzuerkennen, dass die Pseudonymisierung keinen umfassenden Schutz für die betreffenden Daten bieten kann, weil sie Inferenzangriffe nicht vollständig abzuwehren in der Lage ist. Pseudonymisierungs- und Anonymisierungsdienste müssen ihre Praktiken durch Risikobewertungen, Risikominderungsstrategien und Policies zur Einverständniserklärung oder andere Verfahren zur Analyse/Vorverarbeitung/Nachverarbeitung der Daten ergänzen. Die Stelle, bei der die pseudonymisierten Datenbehälter hinterlegt sind, trägt die Verantwortung für die Überprüfung der Datenbehälter auf Inferenzrisiken und für den Schutz vor Offenlegung einzelner Datensatzergebnisse. Die Informationsquelle trägt die Verantwortung für die Vorprüfung/Vorverarbeitung der offengelegten Quelldaten, um die offengelegten Daten vor Inferenz basierend auf Ausreißern, eingebetteten zur Identifizierung geeigneten Daten oder ähnlichen unabsichtlichen Offenlegungen zu schützen.

6.1.2.8 Privatheit und Sicherheit

Es besteht immer das Risiko, dass pseudonymisierte Daten mit der betroffenen Person verknüpft werden können. In Anbetracht dieses Risikos sollten die gesammelten Daten als „personenbezogene Daten“ angesehen und nur für die Zwecke genutzt werden, für die sie erfasst wurden. In vielen Ländern verlangen die geltenden Gesetze, pseudonymisierte Daten in gleicher Weise zu schützen wie zur Identifizierung geeignete Daten.

7 Wiedererkennungsprozess

Es sind zwei verschiedene Kontexte der Wiedererkennung von pseudonymisierten Informationen zu berücksichtigen:

- Wiedererkennung als Teil der normalen Datenverarbeitung;
- Wiedererkennung als außergewöhnliches Ereignis.

Teil der normalen Verfahren

Falls die Wiedererkennung Teil der normalen Datenverarbeitung ist, sollten die für die Wiedererkennung geltenden Bedingungen und die dafür anzuwendenden Verfahren in die Gesamtplanung der Prozesse einbezogen werden. Ein Beispiel ist der Fall, dass pseudonymisierte Anfragen von einer Patientenaktenanwendung in entpersonalisierter Art und Weise an ein klinisches Pathologielabor gesendet werden. Die Ergebnisse werden von der Anwendung in einem pseudonymisierten Format empfangen, dem Wiedererkennungsverfahren unterzogen und automatisch in die Patientenakte eingefügt.

In normalen Verfahren ist die Wiedererkennung dadurch gekennzeichnet, dass sie üblicherweise automatisiert und transparent erfolgt und dass keine Einzelfallberechtigung erforderlich sein sollte.

In Fällen, in denen die Wiedererkennung Teil eines normalen Verfahrens ist, wird auf die Integrität der Daten (Vollständigkeit, keine Änderungen) geachtet. In den meisten dieser Fälle erfordert die Verarbeitung denselben Grad der Integrität wie die personenbezogenen Daten und kann diesen auch garantieren. Dies ist bei Forschungsdaten nicht notwendigerweise der Fall, die aus diesem Grund unter die Kategorie „Ausnahmeverfahren“ fallen.

Ausnahme

Falls die Wiedererkennung eine Ausnahme von der üblichen Datenverarbeitung darstellt, muss der Wiedererkennungsprozess Folgendes erfordern:

- spezielle Authentisierungsverfahren;
- außergewöhnliche Eingriffe durch den Anbieter des Pseudonymisierungsdienstes.

Falls Wiedererkennung von entpersonalisierten Daten als die Ausnahme von der Regel angesehen wird, müssen in der Sicherheits-Policy die Umstände beschrieben werden, die zu einer Wiedererkennung führen können.

In dem für die Sicherheit der Datenverarbeitung geltenden Policy-Dokument sollten die vorhersehbaren Fälle festgelegt werden, und dieses Dokument sollte Folgendes abdecken:

- Jeder Fall sollte beschrieben werden, und es sollten für jeden Fall ein oder mehrere Szenarien für die Wiedererkennung beschrieben werden;
- Identität der Person, die einen Wiedererkennungsauftrag erteilt hat;
- im Abgleich mit den Wiedererkennungsregeln ist zu überprüfen, ob die Person, die den Auftrag erteilt hat, über die entsprechende Berechtigung verfügt. Alle an derartigen Fällen beteiligten Entitäten müssen über das Wiedererkennungsereignis in Kenntnis gesetzt werden. Die beschriebene Wiedererkennung sollte nur bei Vorliegen der entsprechenden Berechtigung (in elektronischer oder anderer Form) gestartet werden und dem in der Policy beschriebenen Szenario folgen;
- eine ausnahmsweise Wiedererkennung sollte nur von einem vertrauenswürdigen Dienstanbieter durchgeführt werden (dies gilt unter der Annahme, dass die Wiedererkennung vom Anbieter des Pseudonymisierungsdienstes auszuführen und dieser dazu in der Lage ist);
- in jedem Falle ist darauf zu achten, dass niemand anders als der vertrauenswürdige Dienstanbieter technisch in der Lage sein darf, Listen zusammenzustellen, die Bezeichner und Pseudonyme miteinander verknüpfen. Nach der Verarbeitung muss der vertrauenswürdige Dienstanbieter diese Verknüpfungslisten zerstören;
- der für die Verarbeitung der re-identifizierten Daten Verantwortliche muss eine gründliche Integritätsprüfung (auf Richtigkeit und Vollständigkeit der Daten) vornehmen. Dies gilt besonders dann, wenn sich der Endzustand der Daten ändert, z.B. bei Umwandlung von pseudonymisierten Forschungsdaten in Daten, die sich für Diagnose- oder Behandlungszwecke nutzen lassen;
- die Policy muss klare Angaben dazu enthalten, wer die Verantwortung für die Verarbeitung der aus dem Wiedererkennungsprozess resultierenden personenbezogenen Daten trägt und in welchem Zustand die Daten letztendlich vorliegen werden. Die wiederhergestellten Daten sollten ihren Ursprung anzeigen (unter Vorbehalt, weil die re-identifizierten Daten unter Umständen nicht mehr vollständig oder nicht so zuverlässig sind wie die ursprünglichen personenbezogenen Daten, von denen sie abgeleitet wurden in Forschungsdatenbanken oder Speichern für klinische Daten).

Die Regeln für vorhersehbare Fälle müssen auch in unvorhersehbaren Ausnahmefällen gültig bleiben. Im Gegensatz zu den vorhersehbaren Fällen gibt es kein *a priori* gültiges Szenario für die Wiedererkennung. Entsprechend muss bewertet werden, wie dringend der Bedarf bezüglich der Wiedererkennung ist. Hierfür trägt der für die Verarbeitung der Daten Verantwortliche die Verantwortung.

Von dieser Regel ausgenommen kann die Wiedererkennung für die Zwecke der Strafverfolgung sein. Dieser Ausnahmefall wird im vorliegenden Dokument nicht behandelt, es wird aber davon ausgegangen, dass die Strafverfolgungsakteure, die die Verantwortung für die Wiedererkennung übernehmen, auch für den vorschriftsmäßigen Schutz der resultierenden personenbezogenen Daten sorgen.

Technische Realisierbarkeit

In Fällen, in denen die Re-Identifizierung Teil des normalen Verfahrens ist oder für eine Anzahl von beschriebenen Szenarien erwartet wird, ist eine Wiedererkennung technisch möglich.

Für die Realisierung der Wiedererkennung stehen verschiedene Verfahren zur Verfügung:

- Zur direkten oder indirekten Identifizierung geeignete Daten (z.B. eine Liste von Ortsbezeichnungen) können gemeinsam mit den pseudonymisierten Daten verschlüsselt und aufbewahrt werden. Nur ein entsprechend benannter vertrauenswürdiger Dienstanbieter kann die Daten entschlüsseln und die zur indirekten Identifizierung geeigneten Daten wieder der betroffenen Person zuordnen.
- Ein vertrauenswürdiger Dienstanbieter (der Anbieter des Pseudonymisierungsdienstes oder ein als Hinterlegungsstelle fungierender Dienstanbieter) kann eine Liste führen, in der Pseudonyme und (zur direkten oder indirekten Identifizierung geeignete) Bezeichner miteinander verknüpft sind.

Anhang A (informativ)

Szenarien für die Pseudonymisierung im Gesundheitswesen

A.1 Einleitung

Dieser Anhang stellt eine Reihe von Fällen oder „Szenarien“ des Gesundheitswesens der höchsten Ebenen vor, die geschäftliche und technische Kernanforderungen an die Pseudonymisierungsdienste darstellen, die einen breiten Querschnitt der Gesundheitswirtschaft unterstützen.

Zuerst werden allgemeine Anforderungen dargelegt, wobei die wesentlichen Grundsätze des Schutzes von personenbezogenen Daten und der Sicherheit sowie die wesentlichen Bedürfnisse der Gesundheitswirtschaft vorgestellt werden. Danach werden die einzelnen Szenarien im vorliegenden Dokument wie folgt ausführlich dargestellt:

- 1) eine Beschreibung des Szenarios oder der Situation des Gesundheitswesens, die gesundheitsbezogene Pseudonymisierungsdienste erfordert;
- 2) daraus resultierende geschäftliche und technische Anforderungen, die ein Pseudonymisierungsdienst erfüllen muss.

A.2 Erläuterung der Szenarien

Die in A.3.1 bis A.3.5 beschriebenen Szenarien zeigen, wie Pseudonymisierungsdienste im Gesundheitswesen genutzt werden können. Jedes Szenario soll mögliche und wahrscheinliche Nutzungen eines gesundheitsbezogenen Pseudonymisierungsdienstes beschreiben.

Bei der Beschreibung der einzelnen Szenarien werden jeweils die folgenden Punkte verwendet:

— **Art der ID**

Darin wird angegeben, ob es sich bei der ID um eine Patienten-ID ist oder z.B. um eine Anbieter-ID handelt.

— **Eindeutigkeit**

In Erwartung des Nutzens, der aus einer pseudonymisierten Datenbank gezogen wird, ist es wichtig, zu wissen, ob der Eingabewert dazu geeignet ist, eine Person in einem gegebenen Kontext eindeutig zu identifizieren. Dies ist besonders wichtig, wenn über einen bestimmten Zeitraum und Organisationsgrenzen hinweg erfasste Daten eindeutig miteinander verknüpft werden sollen. Es ist ebenfalls wichtig, zu bewerten, ob von derselben Entität stammende Daten miteinander verknüpft werden können oder ob das Risiko besteht, dass in der (den) Zieldatenbank(en) Synonyme enthalten sind.

— **Sensitivität der Daten**

Um die Lösung entwerfen zu können, ist es hilfreich, einen Hinweis darauf zu erhalten, wie sensibel die betreffenden Daten sind. Die Sensitivität muss vor dem Hintergrund der geltenden Gesetze oder in Bezug auf die Wichtigkeit im Geschäftszusammenhang bzw. für den Anwendungsfall ausgewertet werden. So hat z.B. die Erfassung von Informationen zu HIV bei natürlichen Personen vom rechtlichen Standpunkt aus einen wesentlichen höheren Sensitivitätsgrad und erfordert eine entsprechende Risikoanalyse. Die Erfassung von Erfolgsquoten einer bestimmten Krankheitsbehandlung bei teilnehmenden Institutionen ist vom rechtlichen Standpunkt aus nicht sensibel, kann dies aber auf dem kritischen Weg einer Geschäftslösung sein oder werden.

- **Datenquellen: singuläre oder multiple Datenquellen und ihre Beziehungen**
Die Anzahl und der Kontext von Datenquellen bestimmen in hohem Maße, ob die Inanspruchnahme der Dienste einer Zwischenorganisation, die vertrauenswürdige Dienste leistet, erforderlich ist oder nicht.
- **Primäre oder indirekte Nutzung von personenbezogenen Daten**
Dies ist eine Eigenschaft, die erheblichen Einfluss auf die rechtlichen Vorgaben hat, die zu unterschiedlichen Entwürfen für die Pseudonymisierungslösung führen können. Es ist auch wichtig, zu wissen, ob die Daten direkt von der betroffenen Person bezogen wurden.
- **Kontext/Endzustand: kommerzielle und medizinische Forschung, Behandlung von Patienten**
Darin wird der Kontext kurz beschrieben.
- **Durchsuchbarkeit/Verknüpfbarkeit**
Die Durchsuchbarkeit ist ein sehr wichtiges Element in der Gesamtplanung einer Pseudonymisierungslösung. Die Granularität der Durchsuchbarkeit muss festgelegt werden, wobei sich die Durchsuchbarkeit auf eine Auswahl von pseudonymisierten Daten basierend auf nicht pseudonymisierten Elementen (z. B. je geographischer Region) bezieht. Die Suchfunktion erfordert die Nutzung eines Pseudonymisierungsdienstes und kann eingeschränkt sein.
- **Umkehrbarkeit/Wiedererkennung**
Dies dient der Berücksichtigung, ob eine Wiedererkennung wünschenswert oder verboten oder unter geregelten Umständen wünschenswert ist oder ob sie für derzeit noch unbekannte, in der Zukunft aber möglicherweise erwünschte Umstände eingebaut werden sollte. Es muss erwogen werden, bis zu welchem Grade eine Wiedererkennung annehmbar ist.
- **Verknüpfung über die Zeit**
Das Wiedererkennungsrisiko wird durch die Menge an pseudonymisierten Informationen beeinflusst, die gesammelt werden kann. Durch die zeitliche Begrenzung der Verknüpfbarkeit lässt sich die Menge an pseudonymisierten Informationen begrenzen. Dies kann natürlich mit den für Langzeitforschungen geltenden Anforderungen kollidieren.
- **Verknüpfung über Bereiche**
Die Anwendung eines bestimmten Pseudonymisierungsschlüssels oder -verfahrens sollte auf einen möglichst engen Bereich begrenzt werden. Deshalb ist es in Bezug auf die Szenarien wichtig, den Bereich zu beschreiben, in dem ein Pseudonym verwendet wird, und anzugeben, für wie lange dieses verwendet wird und welche Art der Verknüpfung mit anderen Bereichen gefordert ist. Dies wiederum bestimmt den Bedarf einer Zwischenorganisation.

Dieser Aspekt kann gegebenenfalls auch der Zusammenarbeit verschiedener Zwischenorganisationen Rechnung tragen.

A.3 Szenarien im Gesundheitswesen

Tabelle A.1 — Eigenschaften der Szenarien

Szenario		Betroffene Person			Datenquellen		Funktions-/Leistungsanforderungen			
		Art der ID	Eindeutigkeit	Sensitivität (Gesetze zu personenbezogenen Daten)	Datenquellen	Primär/indirekt	Kontext/Endzustand	Durchsuchbarkeit	Wiedererkennung	Verknüpfung über die Zeit
1	Pseudonymisierte Pflege	Patienten-ID	Im auflösenden System eindeutig (KIS)	Hoch	Einzelne Datenquelle	Primär	Pflege	n.a.	Ja	Ja
2	Klinische Studie	Patienten-ID	Keine garantierte Eindeutigkeit	Hoch	Mehrere Zentren	Primär	Forschung	Ja	Nein (Ausnahme Policy)	Ja/Nein
3	Klinische Forschung	Patienten-ID/Provider-ID	Keine garantierte Eindeutigkeit	Hoch	Mehrere Zentren	Indirekt	Forschung	Ja	Nein (Ausnahme Policy)	Ja
4	Überwachung der öffentlichen Gesundheit	Patienten-ID/Provider	Keine garantierte Eindeutigkeit	Hoch	Mehrere Zentren	Primär/indirekt	Verwaltung des öffentlichen Gesundheitswesens	Ja	Ja unter sehr genau geregelten Umständen	Ja
5	Berichte zur Patientensicherheit	Patienten-ID/Provider	Eindeutig	Hoch	Mehrere Zentren	Primär	Forschung	Ja	Ja	Ja
6	Nicht das Gesundheitswesen betreffende Forschung	Patienten-ID, andere Bereichs-IDs	Sehr heterogen, keine Eindeutigkeit	Hoch für den Teil der medizinischen Daten	Mehrere Zentren	Indirekt	Nicht-medizinische Forschung	Ja	Nein	Ja
7	Das Gesundheitswesen betreffende Marktforschung	Arzt-ID/Patienten-ID	Eindeutig	Gering (Arzt) Hoch (Diagnose)	Mehrere Zentren	Indirekt	Nicht-medizinische Forschung	Ja	Nein	Ja
8	Lehrdateien	Patienten-ID, andere Bereichs-IDs	Sehr heterogen, keine Eindeutigkeit	Hoch für den Teil der medizinischen Daten	Mehrere Zentren	Primär	Ausbildung	Nein	Nein	Ja
9	Außendienst	Patienten-ID, andere Bereichs-IDs	Sehr heterogen, keine Eindeutigkeit	Hoch für den Teil der medizinischen Daten	Mehrere Zentren	Primär/indirekt	Kommerzielle Betriebe	Nein	Nein	Begrenzt (muss Zeit, nicht aber Verknüpfungen erhalten)

Szenarien

- 1) Pseudonymisierte Pflege (en: Pseudon. Care)
- 2) Klinische Studien und Überwachung nach Markteinführung (en: Clin-trial)
- 3) Indirekte Nutzung von klinischen Daten, z. B. zu Forschungszwecken (en: Clin-res)
- 4) Überwachung und Bewertung der öffentlichen Gesundheit (en: Pub health monitor)
- 5) Vertrauliche Berichte zur Patientensicherheit (en: Patient safety reporting; einschließlich Medikamentennebenwirkungen)
- 6) Nicht das Gesundheitswesen betreffende Forschung (en: Non-HC Research; zuvor Verbrauchergruppen)
- 7) Das Gesundheitswesen betreffende Marktforschung (HC Market Research; einschließlich vergleichender Berichte zu Qualitätsindikatoren, Kreuzgutachten, Auslastung, klinischer Qualifikation/Korrektheit von Arztrechnungen, Finanzabrechnung)
- 8) Lehrdateien (Lehrmaterial, Studienmaterial für Studierende, Spezialfälle für Ärzte)
- 9) Außendienst (muss alle Einzelheiten zu den Maschinen und den mit Hilfe von Maschinen gemessenen Daten bewahren, kann aber üblicherweise allen patienten- und ärztebezogenen Daten und alle Finanzdaten eliminieren)

A.3.1 Reihenfolge nach der klinischen Pathologie (pseudonymisierte Pflege)

Beispielszenarien (in dieser Gruppe):

Dieses Szenario nutzte den Pseudonymisierungsdienst zum Schutz von Patientenidentitäten und zur gleich bleibenden Verfolgung von Patienten über disparate Systeme.

Ein klinischer Leistungserbringer muss eine Probe zur Laborprüfung einsenden. Nach der geltenden Policy dürfen mit dem entsprechenden Auftrag keine zur Identifizierung des Patienten geeigneten Informationen mitgesendet werden. Es ist jedoch wichtig, sowohl die Auftragsbestellung mit dem Auftragsergebnis abzugleichen als auch dafür zu sorgen, dass der Labordienst in der Lage ist, für denselben Patienten ein zeitversetztes Vergleichsergebnis zu liefern. Bevor der Auftrag an das Labor gesendet wird, erstellt ein vertrauenswürdiger Pseudonymisierungsdienst ein Pseudonym, und der Ergebnisdatensatz wird zusammen mit dem Pseudonym ausgegeben. Das Pseudonym wird wiedererkannt, um das Ergebnis in die entsprechende Patientenakte aufzunehmen.

Akteure: Auftragserteiler (z. B. Leistungserbringer im Krankenhauskontext), Auftragserfüller (z. B. klinisches Pathologielabor), Pseudonymisierungsdienst, KIS.

Vorbedingungen: der Auftragserteiler wählt eine Reihe von Tests aus, die der Auftragserfüller auszuführen hat: Der im Rahmen dieses Auftrags erstellte Datensatz wird mit Hilfe einer für das Krankenhaus eindeutigen Kennnummer auf die betroffene Person bezogen.

Nachbedingungen: der Auftragserteiler hat Ergebnisse vom Auftragserfüller erhalten und sie mit Hilfe der für den Auftrag verwendeten und für das Krankenhaus eindeutigen Kennnummer der betroffenen Person in die Patientenakte dieser Person eingetragen.

Ablaufplan/Ereignisse/Maßnahmen

- Erteilung eines Auftrags an das Gesundheitsinformationssystem (KIS):
 - der Auftragserteiler authentisiert gegenüber dem KIS;
 - der Auftragserteiler reicht den Auftrag mit der für das Krankenhaus eindeutigen Kennnummer der betroffenen Person beim KIS ein;
 - der Auftragserteiler überprüft den Auftrag in Bezug auf die geltenden Policies (z. B. der Fall, dass es Empfänger nicht gestattet ist, zur Identifizierung geeignete Daten zu erhalten, VIP usw.) und trifft die Entscheidungen über die zu ergreifenden Datenschutzmaßnahmen.

- Pseudonymisierung:
 - das Krankenhausinformationssystem ruft den Pseudonymisierungsdienst auf mit der für das Krankenhaus eindeutigen Kennnummer als Eingabe;
 - der Pseudonymisierungsdienst verarbeitet die krankenhausspezifischen Kennnummern;
 - der Pseudonymisierungsdienst gibt das Pseudonym an den KIS aus.
- Der KIS sendet den Auftrag mit dem Pseudonym an den Auftragserfüller:
 - Kommunikationsverbindung herstellen;
 - Nachricht gesendet;
 - Bestätigung (en: acknowledge, ack) erhalten.
- Der Auftrag wird vom Auftragserfüller unter Anwendung des Pseudonyms bearbeitet:
 - (mögliche vergleichende Analyse werden vom Spezialisten vorgenommen).
- Der Auftragserfüller legt dem KIS das Ergebnis unter dem betreffenden Pseudonym vor:
 - Kommunikationsverbindung herstellen;
 - Nachricht gesendet;
 - Bestätigung erhalten.
- Wiedererkennung der Ergebnisse:
 - das KIS legt den Pseudonymisierungsdiensten das Pseudonym vor;
 - der berechtigte Nutzer (KIS) wird in Bezug auf die Umkehr-ID-Policy überprüft;
 - der Pseudonymisierungsdienst verarbeitet das Pseudonym;
 - der Pseudonymisierungsdienst sendet die Real-ID an das KIS.
- Das KIS trägt das Ergebnis mit der Krankenhaus-ID in die Patientenakte ein.

Andere Beispiele/Bemerkungen

Online-Beratungsdienste im Internet – an einer Person geleistete Gesundheitsmaßnahme – immer wieder die gleiche Person.

Eine öffentlich bekannte Person stellt sich selbst bei einem Anbieter von gesundheitsbezogenen Dienstleistungen zur klinischen Versorgung bzw. Pflege vor. Mit dem Wunsch sicherzustellen, dass der Klinikaufenthalt und die Nachbehandlung geheim bleiben mögen, beantragt der Patient die durchgehende Verwendung von pseudonymisierten Bezeichnungen.