Pseudonymisierung für Forschungsdatenbanken und Register

TMF Pseudonymisierungsdienst für Medizinische Forschungsnetze

Sebastian Claudius Semler[1], Andreas Lux[2], Wilhelm Dolle[3], Michael Reng [4], Klaus Pommerening[5]
[1] Telematikplattform für Medizinische Forschungsnetze e.V. (TMF), Berlin
[2] Debold & Lux GmbH, Hamburg
[3] interActive Systems GmbH, Berlin
[4] Klinik und Poliklinik für Innere Medizin I der Universität Regensburg
[5] Institut für Medizinische Biometrie, Epidemologie und Informatik der Johannes Gutenberg-Universität Mainz

Was ist Pseudonymisierung?

In medizinischen Einrichtungen werden grundsätzlich personenbezogene Daten der Patienten, im Folgenden als Patientendaten subsumiert, erhoben, gespeichert und verwaltet. Unter diesen Daten sind die einen Patienten **identifizierenden** Daten zu unterscheiden von den sonstigen medizinischen Daten. Hierbei handelt es sich meistens um die klassischen administrativen Daten einer Person wie z. B. Name, Vorname, Geburtsdatum, Geschlecht, Versicherungsnummer. Zuweilen, in bestimmtem klinischem Kontext oder in speziellen Datensammlungen, können aber auch klinische Parameter wie eine sehr seltene Diagnose in Verbindung mit der Wohnregion oder eine bestimmte genetische Konstellation indirekt „Patienten identifizierend" sein.

Datenschutzrechtlich sind der Gebrauch und die Speicherung von klinischen Daten gemeinsam mit den Patienten identifizierenden Daten unbedenklich, so lange ein direkter Behandlungsbezug besteht.

Ohne direkten Behandlungsbezug ist ein Zugriff auf die Patientendaten für Dritte grundsätzlich nicht erlaubt – dies ist eine zentrale Anforderung an die Berechtigungskonzepte und Sicherheitssysteme jeder Softwarelösung im Gesundheitswesen.

Um klinische Daten für Zwecke der Forschung, der Epidemiologie und der Gesundheitsfürsorge nutzen zu können, ist es grundsätzlich erforderlich, dass diese anonymisiert zur Verfügung gestellt werden. Anonymisierung bedeutet, dass eine Zuordnung der Daten zu einer Person technisch und inhaltlich nicht mehr möglich ist – die medizinischen Daten liegen dann entkoppelt von den Patienten identifizierenden Daten vor.

Anonymisierung ist ein gängiges Verfahren, wann immer Daten aus der behandelnden Klink bzw. der behandelnden Praxis in andere Forschungseinrichtungen oder zentrale Forschungsdatenbanken transferiert werden sollen – Diagnosen, Symptome und Daten zum Krankheitsverlauf und zur Therapie können dergestalt gespeichert und verfügbar gemacht werden, ohne dass die sich dahinter verbergende Person zu ermitteln ist. Der technische Exportvorgang ist vergleichsweise simpel – es können schlichtweg die Personen identifizierenden Items fortgelassen bzw. durch Platzhalter ersetzt werden. Verfahren wie diese können für bestimmte Formen der epidemiologischen Patientenregister zufriedenstellend angewendet werden.

Aus wissenschaftlicher und medizinischer Sicht ist das Verfahren der Anonymisierung jedoch in drei grundlegenden Fällen unbefriedigend:

1. wenn Forschungsvorgänge und Behandlungsvorgänge parallel laufen,
2. wenn eine langfristige Beobachtung des Patienten mit entsprechend longitudinaler Fortschreibung der Forschungsdaten gewünscht ist und dies mehrzeitige Datenexportvorgänge oder Behandlungsschritte an unterschiedlichen Institutionen erfordert,
3. wenn sich aus einer Analyse der Forschungsdaten (durch nicht behandelnde Ärzte und Wissenschaftler) mögliche neue und bessere Behandlungsoptionen ergeben können, die man dem Patienten zukommen lassen muss – dies ist ein Anrecht des Patienten, das sich aus der Teilnahme an Studien und Forschungsprojekten ergibt.

Für den ersten und zweiten Fall ist die Fortschreibung des „anonymen" Falls in den Forschungsdatenbanken oder Registern erforderlich. Es muss demnach gewährleistet sein, dass trotz Auslassung Patienten identifizierender Merkmale die Daten immer derselben Person zugeordnet werden können.

Im dritten Fall ist zusätzlich eine Rückermittlung der dahinter stehenden Person anhand eines primär „anonymen" Falls erforderlich.

Autoren: Sebastian Claudius Semler, Andreas Lux, Wilhelm Dolle, Michael Reng, Klaus Pommerening
Titel: Pseudonymisierung für Forschungsdatenbanken und Register
In: Jäckel (Hrsg.) Telemedizinführer Deutschland, Ober-Mörlen, Ausgabe 2005
Seite: 209-214

In diesen Fällen ist eine reine Anonymisierung unzureichend, da weder eine Fallfortschreibung noch eine Re-Identifikation der Person möglich ist. Daher muss bei diesen Anforderungslagen, wie sie bei einer Vielzahl von Forschungsdatenbanken und Registern, aber auch beim Aufbau von elektronischen Patientenakten im Rahmen der „vertikalen Vernetzung“ und integrierten Versorgung gegeben sind, mit einer Pseudonymisierung operiert werden.

Unter Pseudonymisierung versteht man sinngemäß eine eingeschränkte Anonymisierung – ein klarer Personenbezug auch in der longitudinalen Fortschreibung muss gewährleistet sein. Zugleich ist aber im pseudonymisierten Zustand keinerlei direkte Möglichkeit der Identifikation der sich hinter dem Pseudonym verbergenden natürlichen Person gegeben.

Als Definition führt das Bundesdatenschutzgesetz (BDSG) an: „Pseudonymisieren ist das Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren.“ [1]

Eine Verwaltungsvorschrift zur Durchführung des Brandenburgischen Datenschutzgesetzes führt hierzu näher aus [2]:

„Beim Pseudonymisieren (§ 3 Abs. 3 Nr. 2 BbgDSG) werden Einzelangaben über persönliche oder sachliche Verhältnisse einer natürlichen Person auf die Weise ersetzt, dass die verbleibenden Angaben zwar noch einem Einzelnen, nicht mehr jedoch ohne Kenntnis der Zuordnungskriterien einer bestimmten natürlichen Person zugeordnet werden können.

Das Ziel der Pseudonymisierung besteht insbesondere angesichts der zunehmenden Technisierung darin, personenbezogene Daten ohne Kenntnis der Identität des Betroffenen nutzen zu können. Sie sollte überwiegend dort eingesetzt werden, wo die wesentlich stärkere Maßnahme der Anonymisierung personenbezogener Daten nicht möglich ist.

Die Pseudonymisierung kann zum Beispiel im Medizinbereich zum Einsatz kommen. Ein Außenstehender (beispielsweise ein Techniker) kann die Daten keiner Person zuordnen. Ihm gegenüber wird das Arztgeheimnis gewahrt. Das Krankenhauspersonal hingegen kann die Daten ohne Einschränkung verwenden.“

Relevante Definitionen aus dem Bundesdatenschutzgesetz (BDSG), Fassung vom 14.01.2003

§ 3 Weitere Begriffsbestimmungen

(1) **Personenbezogene Daten** sind Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbaren natürlichen Person (Betroffener).

(6) **Anonymisieren** ist das Verändern personenbezogener Daten derart, dass die Einzelangaben über persönliche oder sachliche Verhältnisse nicht mehr oder nur mit einem unverhältnismäßig großen Aufwand an Zeit, Kosten und Arbeitskraft einer bestimmten oder bestimmbaren natürlichen Person zugeordnet werden können.

(6a) **Pseudonymisieren** ist das Ersetzen des Namens und anderer Identifikationsmerkmale durch ein Kennzeichen zu dem Zweck, die Bestimmung des Betroffenen auszuschließen oder wesentlich zu erschweren.

(9) **Besondere Arten personenbezogener Daten** sind Angaben über die rassische und ethnische Herkunft, politische Meinungen, religiöse oder philosophische Überzeugungen, Gewerkschaftszugehörigkeit, **Gesundheit** oder Sexualleben.

§ 3a Datenvermeidung und Datensparsamkeit

Gestaltung und Auswahl von Datenverarbeitungssystemen haben sich an dem Ziel auszurichten, keine oder so wenig personenbezogene Daten wie möglich zu erheben, zu verarbeiten oder zu nutzen. **Insbesondere ist von den Möglichkeiten der Anonymisierung und Pseudonymisierung Gebrauch zu machen**, soweit dies möglich ist und der Aufwand in einem angemessenen Verhältnis zu dem angestrebten Schutzzweck steht.

(Hervorhebungen durch die Autoren.)

Abbildung 1: *Begriffsdefinitionen im BDSG*

Zusammenfassend lässt sich unterscheiden [3, 4]:

- Bei Personenbezug ist das einzelne Individuum in seiner Identität erkennbar.
- Bei Anonymität besteht ein Bezug nur auf eine Gesamtmenge von Individuen.
- Bei Pseudonymität besteht ein Bezug zu einem einzelnen Individuum, dessen Identität allerdings nicht erkennbar ist.

Durch die Pseudonymisierung „kann der Personenbezug so verschleiert werden, dass faktische Anonymität entsteht, ohne die gewünschte Verwertung der Daten nennenswert zu behindern“ [4].

Die Etablierung einer Lösung zur Pseudonymisierung und zur optionalen Depseudonymisierung unter Einhaltung aller Datenschutz rechtlichen Erfordernisse ist eine wesentliche Herausforderung für alle medizinischen Forschungsnetze, zugleich eine wichtige Voraussetzung für den Aufbau und Betrieb von Patientenregistern und Forschungsdatenbanken.

Konsentierte „Generische Datenschutzkonzepte“ und Pseudonymisierung für Forschungsdatenbanken und Register

Als gemeinsame, übergreifende Aufgabe wurde diese Herausforderung 2001 durch die vom Bundesministerium für Bildung und Forschung (BMBF) geförderten Kompetenznetze in der Medizin unter dem Dach der Telematikplattform für Medizinische Forschungsnetze (TMF) angegangen [5, 6].

Die TMF ist eine ebenfalls BMBF-geförderte Plattform für die vernetzte klinische Forschung – als Interessengemeinschaft der medizinischen Kompetenznetze (KN), der Koordinierungszentren für klinische Studien (KKS) und weiterer vom BMBF geförderter medizinischer Forschungsverbünde dient sie der Entwicklung und dem Aufbau von nachhaltigen IT-Infrastrukturen in der medizinischen Forschung sowie ganz allgemein der effektiveren Gestaltung der Rahmenbedingungen für die medizinische Forschung in Deutschland [7, 8].

Innerhalb der TMF hat sich die Arbeitsgruppe Datenschutz und Datensicherheit der Forschungsverbünde erfolgreich dem datenschutzrechtlichen Rahmen gewidmet. Es konnte ein „Generisches Datenschutzkonzept" formuliert, mit dem Arbeitskreis Wissenschaft der Landesdatenschützer 2003 konsentiert und veröffentlicht werden [9, 10, 11]. Dieses unterscheidet unterschiedliche Szenarien, je nach Ausrichtung der Forschungsnetze, für:

- klinisch-wissenschaftlich fokussierte Forschungsnetze, bei denen vor allem Ärzte und ihre Mitarbeiter aus dem klinischen Behandlungsprozess heraus an der Erhebung und Dokumentation der Daten beteiligt sind,

und

- wissenschaftlich-klinisch fokussierte Forschungsnetze, bei denen die Datenerhebung durch Wissenschaftler abseits des klinischen Behandlungsprozesses erfolgt. Hier steht nicht der Komplex der Erhebung und Nutzung von Daten im Behandlungszusammenhang, sondern die Prozessierung von Daten, die für die wissenschaftliche Forschung bereitgestellt werden, im Fokus.

Der erfolgte bundesweite Abstimmungsprozess mit den Landesbeauftragten hat zur Folge, dass die Konzepte von anderen Forschungsverbünden in identischer oder leicht abgewandelter Form übernommen werden können und dass dabei eine Verkürzung der datenschutzrechtlichen Einzelbegutachtung erreicht wird. Das Konzept für klinisch-wissenschaftliche Forschungsnetze wird hier nicht im Detail beschrieben [vgl. 9-11].

Im generischen Datenschutzkonzept für wissenschaftlich-klinisch fokussierte Forschungsnetze werden drei Grundforderungen dargelegt:

- Es bedarf eines Verfahrens zur sicheren Identifikation jedes Patienten als Individuum, ohne eine klartextliche Kenntlichmachung der Person erforderlich zu machen. Das soll mittels eindeutiger, nicht sprechender Zeichenketten als Patientenidentifikatoren (PID) erfolgen, bei deren Erzeugung die Bildung von Synonymen und Homonymen vermieden werden muss. (Dies würde zu falschen distinkten virtuellen Individuen im pseudonymisierten Datenbestand führen. Stattdessen ist eine sichere Zusammenführung zu einem pseudonymisierten Fall erforderlich.)
- Es bedarf einer Option, für die Forschungsdaten eine Qualitätssicherung durchzuführen, um Plausibilität und Vollständigkeit so weit wie möglich zu sichern.
- Für die Pseudonymisierung der Forschungsdaten soll ein Verfahren hoher Sicherheit eingesetzt werden; die Schlüssel zur Aufdeckung von Pseudonymen dürfen nur an zentraler Stelle treuhänderisch verwaltet werden.

Daraus abgeleitet werden folgende technisch-organisatorische Instrumente für eine entsprechende IT-Lösung gefordert: [9, 10]

- die sog. Patientenliste (mit der Funktion eines PID-Generators)
- der Pseudonymisierungsdienst
- die Möglichkeit der Qualitätssicherung der Erhebungsdaten hinsichtlich Vollständigkeit und Plausibilität
- Protokolle und kryptografischen Verfahren für die Sicherung der Datenübermittlung („kanalorientierte" Sicherheit) und für die Dokumentensicherheit

Ziel dieses als Dienst in Netzen zur Verfügung zu stellenden Systems soll es sein, patientenbezogen medizinische Informationen – je nach Fragestellung in unterschiedlichem Ausmaß – zu erfassen und in einer zentralen Datenbank pseudonymisiert zu speichern.

Konzept und Komponenten des TMF-Pseudonymisierungsdienstes

Basierend auf diesen funktionellen und Datenschutz rechtlichen Rahmenvorgaben wurde mangels verfügbarer Alternative auf dem Anbietermarkt ein innerhalb der TMF gefördertes Projekt gestartet zur Schaffung einer solchen einsatzfähigen Lösung für die Forschungsverbünde.

Die Serverkomponente **„Patientenliste"** dient der möglichst eindeutigen Identifikation eines Patienten anhand eines nicht sprechenden Identifikators. Das gewählte Verfahren ist dabei in der Lage, auch Abweichungen in den klartextlichen Eingabedaten bei mehrfacher Meldung im Zeitverlauf abzufangen – eine Toleranz gegenüber Schreibfehlern (Buchstabendreher, Umlaute etc.) ist damit gewährleistet.

An dieser „Patientenliste" wird der – mit einer rechtsgültigen Einwilligung – an einer Studie teilnehmende Patient von durchaus verschiedenen Einrichtungen und zu unterschiedlichen Zeitpunkten angemeldet. Technisch betrachtet, werden eingegebene oder aus einer Krankenversichertenkarte (KVK) ausgelesene klartextuelle Personen identifizierende Daten wie Name, Geburtsdatum etc., im Folgenden als IDAT bezeichnet, von dem anmeldenden EDV-System an die „Patientenliste" übermittelt. Die „Patientenliste" prüft daraufhin anhand komplexer Algorithmen zunächst im Bestand auf gleiche oder ähnliche Schreibweisen der IDAT, generiert, wenn keine Übereinstimmung vorliegt, eine PID (**PID-Generator**-Funktion) und übermittelt schließlich die IDAT zusammen mit der im Bestand identifizierten oder neu erzeugten PID zurück an die meldende Klinik.

Gerade bei chronischen oder rezidivierenden Erkrankungen ist es wichtig, dass Fallverläufe möglichst langfristig verfolgt werden können. Doch eben bei diesen Verläufen ist die Wahrscheinlichkeit hoch, dass über die Zeit – durch Wechsel von Behandlungseinrichtungen oder räumliche Mobilität der Patienten – die Patienten von verschiedenen Einrichtungen an das Register gemeldet werden. Das vorliegende System ist daher so ausgerichtet,

dass auch bei modifizierter, fehlerbehafteter Eingabe der IDAT durch den Fehlertoleranz gewährleistenden Algorithmus ein Patient im Bestand dennoch eindeutig identifiziert und ihm die identische PID zugewiesen wird [12].

Zu der dergestalt erzeugten eindeutigen, Personen identifizierenden PID können nunmehr in den Datenerfassungssystemen der behandelnden bzw. Studien betreuenden Einrichtungen die zu erfassenden medizinischen Daten [MDAT] patientenbezogen zugeordnet werden.

Nach Abschluss des Erfassungsintervalls, Qualitätssicherung und Freigabe der Daten aus den Erfassungssystemen können nunmehr die MDAT (in verschlüsselter Form als $(MDAT)_{Cr}$) und die PID an die zentrale Forschungsdatenbank übergeben werden. Hier jedoch darf die zunächst im Behandlungszusammenhang bekannte PID nicht das Personen identifizierende Kriterium sein – daher muss der Datensatz zunächst pseudonymisiert werden.

Der eigentliche **Pseudonymisierungsdienst** (PSD) oder **Pseudonymisierungsserver** leistet die Pseudonymisierung derjenigen medizinischen Daten, die außerhalb der Behandlungseinrichtung für Forschungszwecke bereitgestellt werden sollen. Technisch besteht der Pseudonymisierungsvorgang in der kryptografischen Transformation der PID in ein Pseudonym (PSN). Hierbei wird ein symmetrischer kryptografischer Algorithmus hoher Sicherheit genutzt (3DES). Der Schlüssel ist (gegen Auslesen gesichert) auf einer am Server in einem Kartenleser befindlichen Chipkarte gespeichert. Die Transformation des PID in ein PSN wird auf dieser Chipkarte durchgeführt, so dass der geheime Schlüssel die Karte nicht verlässt. Die Pseudonymisierung ist demnach ein rein serverbasierter Vorgang, eine „Maschinenfunktion“, die keiner Benutzerinteraktion bedarf. Das Geheimnis des PSN wird einerseits durch die Art des Workflow, andererseits durch die Verwahrung der Schlüsselkarte sichergestellt.

Die entsprechende Chipkarte wird einmal von der „Certification Authority“ in einem datenschutzbegleiteten Vorgang erstellt. Sicherheitsduplikate sind in sicherer treuhänderischer Verwahrung zu verwalten. Wichtig ist: An keiner Stelle dürfen PID und PSN oder gar PID, klartextuelle Patientenidentifikatoren und PSN gemeinsam bekannt sein. Daher sind auch die Server der „Patientenliste“ und des Pseudonymisierungsdienstes getrennt von einander und an jeweils zentraler Stelle zu betreiben.

Eine analoge Umschlüsselung der Arzt- und Klinikidentifikatoren wird optional angeboten.

Das Pseudonym (PSN) kann nun mitsamt der weiterhin „gekrypteten“, vom Pseudonymisierungsvorgang völlig unbetroffenen MDAT an die jeweilige Forschungsdatenbank weitergeleitet werden. Dort dient das PSN als Zuordnungskriterium für Speicherung und Zusammenführung der Daten und für alle fallbezogenen Auswertungen, die daraus abgeleitet werden.

Falls diese fallbezogenen Auswertungen den Bedarf ergeben, Kontakt zum Patienten aufzunehmen – sei es zu einer ethisch erforderlichen Veränderung der Therapie, sei es im Rahmen von Rekrutierungsanfragen zur Teilnahme an einer klinischen Studie – so kann ein **Depseudonymisierungsverfahren** in Gang gesetzt werden. Auch könnte ein Patient sein Auskunftsrecht über die in der Forschungsdatenbank über ihn (pseudonymisiert) gespeicherten Daten geltend machen.

Hierzu bietet der Pseudonymisierungsserver die Funktion, einen „Depseudonymisierungsantrag“ entgegen zu nehmen. Technisch betrachtet bekommt der PSD eine Anfrage mit dem jeweils betroffenen PSN (mit oder ohne MDAT) gesendet. Diese bedarf allerdings in jedem Fall der Benutzerinteraktion im Sinne einer Freigabe des weiteren Verfahrens. Erfolgt diese, so kann über den oben beschriebenen, SmartCard basierten Algorithmus eine Rückberechnung vom PSN zur PID erfolgen mitsamt Weiterleitung der Anfrage und der weiteren Daten an die jeweilige letzte Behandlungs- bzw. Betreuungsstelle.

Hierzu ist organisatorisch sicherzustellen, dass dieses Freigabeverfahren über eine lokale Ethik- oder Datenschutzkommission überwacht und ausgelöst und in geeigneten Regelwerken auditierbar festgeschrieben wird, da anderenfalls Antrag und Bewilligung einer Depseudonymisierung jeweils im Einzelfall der datenschutzrechtlichen Prüfung und Genehmigung bedürfen.

Die Kommunikationsstrecken zwischen den verteilt im Netz operierenden Komponenten müssen ebenfalls gesichert sein. So ist im System für die Kanal orientierte Sicherheit die Verwendung von Secure Sockets Layer (SSL) mit einer Schlüssellänge von 1.024 Bit vorgesehen. Diese ist zumindest für jede Datenübertragung außerhalb von lokalen Netzen einzelner Krankenhäuser vorgesehen.

Weiterhin sind für die Server zur Authentifizierung Software-basierte Schlüssel und Zertifikate, für Teilnehmer in den Kliniken und Praxen zur Authentifizierung und Dokumentenverschlüsselung Chipkarten gemäß neuer HPC-Spezifikation vorgesehen. Alternativ kann mit softwarebasierten Schlüsseln und Zertifikaten für die Nutzer gearbeitet werden. Eine entsprechende PKI ist (Teilnehmerservice, Trustcenter) ist für jede Installation bereitzustellen.

Da alle Komponenten im Intra- und Internet (wie dargestellt, über sichere Verbindungen) kommunizieren und lokale Installationen weitgehend vermieden werden sollen, sind alle Frontends vollständig webbasiert spezifiziert und umgesetzt. Die Funktionen laufen in gängigen Webbrowsern. Hierfür ist ein spezieller „Security-Proxy“ entwickelt worden, der alle Hardwarezugriffe auf die Security-Komponenten und PKI-Abfragen kapselt.

Schließlich sind für den Testaufbau des Gesamtsystems exemplarische Komponenten zur Patientenanmeldung und lokalen Datenerhebung („Klinikkomponente“) und als zentrale Forschungsdatenbank entwickelt worden.

Das Projekt der TMF

Im Auftrag des Koordinierungsrates des TMF und der AG Datenschutz und Datensicherheit der TMF wurde unter der Federführung des Fraunhofer-Instituts für Software- und Systemtechnik (ISST) und der Debold & Lux GmbH gemeinsam mit mehreren Partnern Ende 2001 begonnen, Spezifikationen für die Einzelkomponenten eines solchen Systems zur Pseudonymisierung zu erstellen und umzusetzen.

Nachdem die Prototypen der Einzelkomponenten Ende 2003 vorlagen, wurden in neuen Förderprojekten der TMF unter der Projektleitung der TMF-

Geschäftsstelle die Einzelkomponenten gesichtet, zum Teil aktualisiert und dokumentiert. Dies geschah unter Beteiligung der Partner

- Debold & Lux GmbH, Hamburg
- interActive Systems (iAS), Berlin
- Institut für Medizinische Biometrie, Epidemologie und Informatik der Johannes Gutenberg-Universität Mainz
- Fraunhofer-Institut Sichere Telekooperation (SIT), Darmstadt
- Trustcenter Schlumberger Sema Competence Center für Informatik (CCI), jetzt zu ATOS Origin gehörig.

Gemeinsam wurden im Laufe des zweiten Quartals 2004, unter Federführung der hierzu beauftragten iAS GmbH die Komponenten in einer Laborinstallation zu einem lauffähigen kommunizierenden Gesamtsystem aus verteilten Softwarekomponenten zusammengesetzt. Mehrere Testinstallationen zu Zwecken von Entwicklung, Test und Support konnten an verschiedenen Standorten etabliert werden.

Nachdem einzelne Komponenten bereits seit 2003 in mehreren Forschungsnetzen im Echtbetrieb laufen, beginnt nun im Laufe der zweiten Jahreshälfte 2004 der Pilotbetrieb des Gesamtsystems zur Pseudonymisierung. Für diese Pilotphase steht eine weitere funktionelle und technische Ausrundung wie auch eine Optimierung der Modularisierung und der Schnittstellen der einzelnen Komponenten auf dem Projektplan.

Für die nahe Zukunft ist die Ansteuerung der elektronischen Gesundheitskarte (eGK) zum Auslesen der Patientendaten, die Implementierung von CDISC in XML-basierten Komponentenschnittstellen und die Integration in verschiedene Remote Data Entry (RDE)-Systeme geplant. Vor allem aber ist die Verstetigung der Lösung mit industriellen Partnern und Betreibern in Vorbereitung.

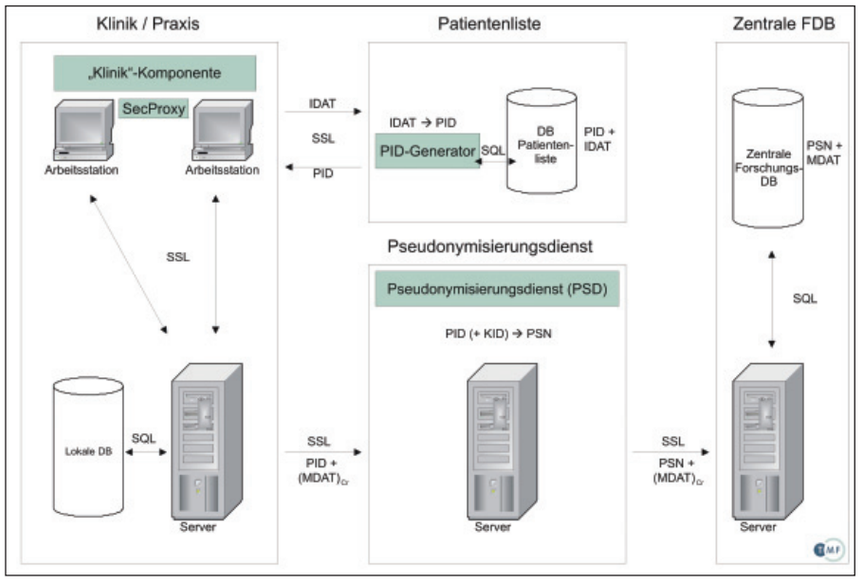

Abbildung 2: *Systemdiagramm Pseudonymisierungsdienst*

Fazit und Ausblick

Die realisierte Komponenten basierte, webfähige und in ihrer Sicherheitsstruktur bereits HPC-konforme Lösung zur Pseudonymisierung stellt einen wichtigen Baustein dar zum Aufbau und zum Betrieb datenschutzkonformer IT-Strukturen für die medizinische Wissenschaft. Sie bietet die Möglichkeit, Patientendaten z. B. für Langzeit- und Querschnittstudien, Statistiken und andere Forschungszwecke zu nutzen, ohne dass die Identität und die Persönlichkeitsrechte des Patienten gefährdet würden.

Ebenso wie die im Rahmen der AG Datenschutz der TMF gemeinsam mit dem AK Wissenschaft der Landesdatenschützer erarbeiteten generischen Datenschutzlösungen könnte sie nachhaltig helfen, den Informationsfluss von der Klinik zur Wissenschaft und aus der Wissenschaft in die Regelversorgung zu optimieren. Wissenschaftliche Errungenschaften könnten somit den Betroffenen rascher und unmittelbarer als bisher zu Gute kommen.

Schließlich kommt man mit schnittstellenfähigen, datenschutzkonformen Lösungen dieser Art der Vision näher, Primärdaten aus im Rahmen der Patientenversorgung anfallenden medizinischen Dokumentation direkt zu überführen und zu nutzen für wissenschaftliche Zwecke, gewissermaßen die wissenschaftliche Dokumentation aus der medizinischen zu generieren. Hierbei könnten aufwendige Doppeldokumentationen vermieden und im Behandlungsprozess etablierte Schritte zur Qualitätssicherung der klinischen Dokumentation auch auf die wissenschaftliche Datenerhebung ausgedehnt werden [13] – je integraler Behandlungsablauf, Dokumentation und wissenschaftliche Auswertung ineinander greifen, desto höherer Mehrwert entsteht.

Die technischen und datenschutzrechtlichen Rahmenbedingungen sind schließlich hinsichtlich der erforderlichen Pseudonymisierung bzw. getrennten Haltung von Patienten identifizierenden und medizinischen Daten ähnlich – sowohl für den Betrieb von Forschungsdatenbanken und Registern, wie auch für longitudinal fortzuschreibende, suprainstitutionelle, nicht durchgängig in einem Behandlungszusammenhang stehende Elektronische Patientenakten.

Ausgewählte Literatur und Verweise

(1) Bundesdatenschutzgesetz (BDSG) in der Fassung vom 14.01.2003
(2) Verwaltungsvorschrift des Ministeriums des Innern zur Durchführung des Brandenburgischen Datenschutzgesetzes (VV zum BbgDSG) vom 22.01.2003
(3) Pommerening K: Pseudonyme - ein Kompromiß zwischen Anonymisierung und Personenbezug. In: Trampisch HJ, Lange S (Hrsg.), Medizinische Forschung - Ärztli-

ches Handeln, 40. Jahrestagung der GMDS, Bochum, September 1995, MMV Medizin Verlag, München 1995, 329-333.
(4) Pommerening K: Chipkarten und Pseudonyme. F!FF Kommunikation 1/96, 9-12
(5) Pommerening K, Wagner M: Ein Pseudonymisierungsdienst für medizinische Forschungsnetze. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 32 (2001), 251
(6) Claßen B: Datenschutz und Datensicherheit. Management & Krankenhaus 12/2002, S.16
(7) Glitscher W, Claßen B: Telematikplattform für medizinische Forschungsnetze der Gesundheitsforschung des BMBF (TMF) – Informationstechnologie für die medizinische Forschung. In: Jäckel, A. (Hrsg.): Telemedizinführer 2003, Medizin Forum Verlag, Ober Mörlen 2002, S. 210-211
(8) Rienhoff O: Bedeutung der Kompetenznetze für die Innere Medizin. Medizinische Klinik 99 [7] (2004). S. 407-411
(9) Debold P, Reng M: Generische Lösungen der TMF zum Datenschutz in der Medizin (2003)
(10) Reng CM, Debold P, Adelhard K, Pommerening K: Generisches Datenschutzkonzept für Forschungsnetze in der Medizin, Shaker-Verlag, Aachen 2004
(11) Reng CM, Debold P, Adelhard K, Pommerening K: Vernetzte medizinische Forschung – Akzeptiertes Datenschutzkonzept. Dt. Ärzteblatt 100 [33] (2003), S. A 2134–2137
(12) Pommerening K, Faldum A: Sicherheit bei der Patientenidentifikation in medizinischen Forschungsnetzen. Informatik, Biometrie und Epidemiologie in Medizin und Biologie 33 (2002), 328
(13) Pommerening K, Reng M: Secondary use of the Electronic Health Record via pseudonymisation. In: Bos L, Laxminarayan S, Marsh A (Eds.): Medical Care Compunetics 1, IOS Press, Amsterdam 2004; pp. 441 – 446
(14) Telematikplattform für Medizinische Forschungsnetze e.V. (TMF): www.tmf-ev.de.

Kontakt
Sebastian Claudius Semler
Wissenschaftlicher Geschäftsführer
Telematikplattform für
Medizinische Forschungsnetze e.V.
(TMF)
Neustädtische Kirchstr. 6
10177 Berlin
sebastian.semler@tmf-ev.de
www.tmf-ev.de