BERECHTIGUNGSKONZEPT FÜR EIN E-HEALTH-PORTAL

Stingl C[1], Slamanig D[1]

Kurzfassung

Zentrale eHealth-Portale werden das Gesundheits-wesen durch einen orts- und zeitunabhängigen Zugriff auf Gesundheitsdaten nachhaltig verbessern. Nichtsdestotrotz müssen Bereiche wie die Kommunikationsinfrastruktur, die Interoperabilität der Subsysteme und speziell der Datenschutz analysiert und verbessert werden. Im Bereich des Datenschutzes wird im Allgemeinen eine Vielzahl von Angriffsszenarien berücksichtigt, jedoch findet die „erzwungene Offenlegung" von Gesundheitsdaten kaum Berücksichtigung. Gerade dieser Aspekt wird in der vorliegenden Arbeit eingehend beleuchtet und Lösungsansätze aufgezeigt.

1. Einleitung

Stehen Gesundheitsdaten in elektronischer Form orts- und zeitunabhängig zur Verfügung, so kann die Effizienz, die Effektivität und die Qualität der medizinischen Behandlung für Patienten enorm gesteigert werden. In vielen Ländern werden derzeit Strategien zur Implementierung einer Elektronischen Gesundheitsakte (EGA) erarbeitet. Dabei ist ein wesentlicher Punkt die Einhaltung der Datenschutzgesetze bzw. Gewährleistung der Datensicherheit. Im Speziellen muss einerseits das einfache Management (inklusive Datenübermittlung) von Patientendaten und andererseits die absolute Geheimhaltung dieser realisiert werden. In Österreich wird ein Zugriffskonzept über einen zentralen Metadatenindex auf dezentral gespeicherte Patientendaten angestrebt [1]. Ein weiterer Grundsatz für die Konzeption einer EGA ist, dass Patienten im Sinne der informationellen Selbstbestimmung über die Preisgabe und Verwendung ihrer Daten selbst entscheiden können. Nichtsdestotrotz birgt die ortsunabhängige Verfügbarkeit von personenbezogenen Gesundheitsdaten ein erhöhtes Potential für Missbrauch. Dies kann die Offenlegung von Gesundheitsdaten, Datendiebstahl, Datenmanipulation und statistische Analysen (z.B. Verknüpfung zwischen Ärzten und Patienten), etc. beinhalten [2].

Eine weitere Problematik, die in dieser Arbeit besondere Betrachtung findet und bis jetzt kaum diskutiert wurde, ist die erzwungene Offenlegung von Gesundheitsdaten (z.B. bei einem Vorstellungsgespräch), die aufgrund der zentralen Zugriffsmöglichkeit ein hohes Gefahrenpotential mit sich bringt. Die hier vorgeschlagene Lösung ermöglicht es einem Patienten, seine Gesundheitsdaten mit mehreren unabhängigen Teil-Identitäten, deren Inhalt jedoch nicht zwingenderweise disjunkt sein muss, zu verwalten. Durch die Definition von sogenannten Rollenpseudonymen kann ein Patient gegenüber verschiedenen Personen (z.B. Ärzten) unterschiedliche Ausschnitte seiner EGA präsentieren und dadurch die unerwünschte Offenlegung kompromittierender Informationen unterbinden. Weiters kann aus der Kenntnis einer Teil-Identität bzw. aus dem Inhalt dieser, nicht auf die Exis-

1 Fachbereich für Medizinische Informationstechnik, FH Technikum Kärnten, Klagenfurt

tenz weiterer Identitäten geschlossen werden. Damit kann ein Patient glaubwürdig weitere Teil-Identitäten abstreiten.

2. Methoden

Als Basis für das Berechtigungskonzept wird ein eHealth-Portal (zentrale Zugriffskomponente) herangezogen, das den Zugriff auf dezentral gespeicherte Dokumente über einen Metadatenindex (Referenzen auf die eigentlichen Dokumente) gewährleistet. Diese Dokumente können optional einen direkten Bezug zum Patienten beinhalten. Das System enthält zudem ein öffentlich zugängliches Verzeichnis aller involvierten Parteien, wie beispielsweise Patienten und Gesundheitsdiensteanbieter (GDA), um eine effiziente Vergabe von Freigaben für Gesundheitsdaten (Dokumente) zu ermöglichen.

Das Ziel dieser Arbeit ist die Definition eines einheitlichen Berechtigungskonzeptes, das auf dem Metadatenindex für Dokumente aufsetzt und die Zugriffe für diese regelt. Die Berechtigungen bzw. Freigaben werden dabei zentral verwaltet, könnten grundsätzlich jedoch auch dezentral (z.B. auf Chipkarten) gespeichert werden. Im Prozess der Freigabenerstellung sind sowohl Gruppen von Personen (Krankenhaus, Abteilung, Station, etc.) als auch Einzelpersonen (Arzt, Patient, Teil-Identität) einbezogen.

Vorerst wird der Begriff einer Teil-Identität erläutert. Jeder Patient kann eine Menge $\mathcal{B}=\{B_1,\ldots,B_n\}$ von Bezeichnern wählen. Diesen kann nachfolgend eine Teilmenge aus der Menge seiner Gesundheitsdaten D, zugeordnet werden. Damit ist eine Teil-Identität I ein 2-Tupel (B_i, D_i) der Relation $\mathcal{B} \times \mathcal{P}(D)$, wobei D_i und D_j für $i \neq j$ nicht disjunkt sein müssen. Abbildung 1 stellt den Zusammenhang zwischen einem Patienten und dessen Teil-Identitäten grafisch dar. Zusätzlich kann in jeder Identität eine Ordnerstruktur definiert werden, um die Gesundheitsdaten besser strukturieren zu können.

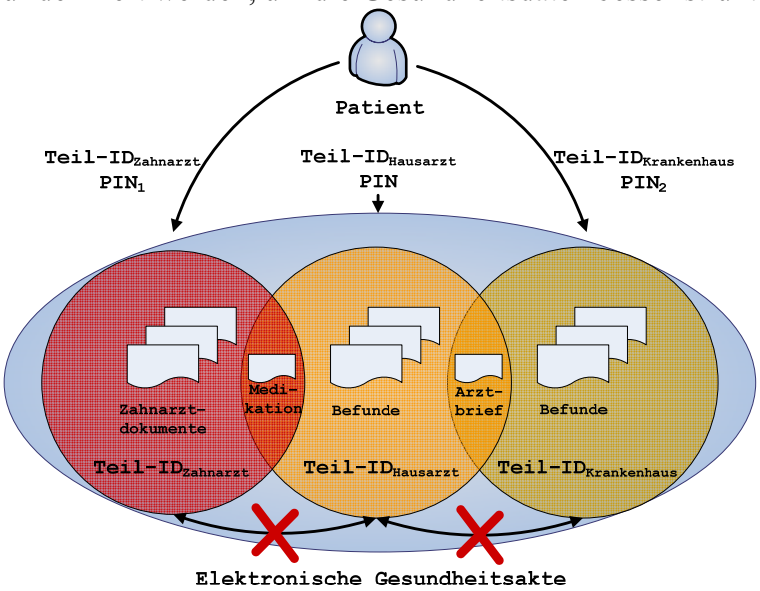

Abbildung 1: Teil-Identitäten eines Patienten mit der öffentlichen Identität (Bezeichner „Hausarzt“).

Jeder Patient besitzt im System genau eine öffentliche Identität (public Identity) und beliebig viele private Identitäten (private Identities). Grundsätzlich werden dem Patienten bei dessen Initialisierung im System die öffentliche Identität und eine feste Anzahl an privaten Identitäten, die standardmäßig inaktiv sind, zugewiesen. Der Patient kann nachfolgend jede inaktive Identität in eine aktive umwandeln. Im System selbst sind aktive und inaktive Identitäten nicht zu unterscheiden (z.B. für einen Administrator). Damit der Patient Zugriff auf seine Gesundheitsdaten erlangt, ist eine erste Authentifikation notwendig. Nach erfolgreicher Authentifizierung wird dem Patienten die

öffentliche Identität zugewiesen und dieser erlangt damit Zugriff auf die damit verknüpften Daten. Dabei ist zu beachten, dass diese Identität keine kompromittierenden Daten beinhalten sollte. Um eine weitere Identität einblenden zu können, muss der Patient für diese Identität eine zusätzliche Authentifikation vornehmen. Mittels der geheimen Authentifikationsdaten (z.B. PIN) wird automatisch die entsprechende Identität eruiert und freigeschaltet. Aus diesem Grund sollten die verwendeten PINs unterschiedlich gewählt werden. Um diese mehrfachen Authentifizierungen praktikabel gestalten zu können, werden folgende Verfahren betrachtet:

1.) Ein Passwort für die erste Authentifizierung und eine PIN pro Identität.
2.) Eine Chipkarte und ein PIN für die erste Authentifizierung und eine weitere PIN pro Identität.
3.) Je eine Chipkarte und ein PIN pro Identität (öffentliche und private).

Aufgrund von Sicherheits- und Usability-Aspekten ist die zweite Variante zu bevorzugen.
Eine weitere wesentliche Eigenschaft dieses Konzeptes ist, dass die Teil-Identitäten eines Patienten nicht eruierbar sind, dass keine Verknüpfung zwischen diesen hergestellt werden kann und auch ein Patient, der gerade eine Teil-Identität verwendet, keine Informationen über weitere Teil-Identitäten ermitteln kann. Ohne diese Eigenschaft wäre, wenn man mit einer Teil-Identität angemeldet ist, ein glaubwürdiges Abstreiten der Existenz einer weiteren nicht möglich.

Die zuvor beschriebenen Einzelpersonen und Personengruppen können bei der Vergabe von Freigaben in den Rollen *C* (Creator), *P* (Patient, Owner), *S* (Sender, Grantor) und *R* (Receiver, Grantee) auftreten (siehe Abbildung 2). Dabei muss beachtet werden, dass es sich hierbei nicht um ein physikalisches „senden“ von Informationen einer Person an eine andere Person handelt, sondern um die Erstellung einer Freigabe für ein Dokument im System.

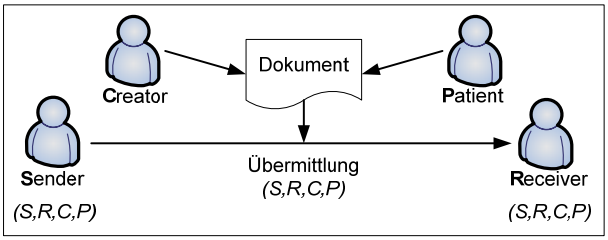

Abbildung 2: Rollen im Freigabeprozess

Eine Referenz auf ein Dokument (im Metadatenindex) kann im Berechtigungssystem in Form eines 4-Tupels *(S,R,C,P)* dargestellt werden. Dabei kann sich das 4-Tupel des Senders von dem des Empfängers unterscheiden. Auf der einen Seite dient das 4-Tupel des Senders zur Protokollierung seiner getätigten Freigaben und auf der anderen Seite zur Protokollierung der empfangenen Dokumentreferenzen. Das 4-Tupel beim Sender wird ausschließlich durch diesen definiert, wohingegen der Sender gezielt Informationen im übermittelten 4-Tupel für den Empfänger ausblenden kann. Der Empfänger selbst kann auf Basis dieses 4-Tupels noch weitere Informationen entfernen. Beispielsweise wird beim klassischen Datenaustausch zwischen GDAs sowohl beim Sender, als auch beim Empfänger das vollständige 4-Tupel gespeichert. Andererseits wird bei der Freigabe von Gesundheitsdaten von einer Teil-Identität an eine andere beim Sender kein 4-Tupel gespeichert und der Empfänger erhält vom Sender das 4-Tupel ohne die Senderinformation *(_,R,C,P)*. Dies bedeutet, dass der Sender aus dem 4-Tupel entfernt wurde und somit keine Verbindung zwischen diesen Teil-Identitäten hergestellt werden kann.

In weiterer Folge werden die wesentlichen Szenarien für die Weitergabe von Berechtigungen zwischen den involvierten Parteien (*G*DAs, *P*atienten, *I*dentitäten) beleuchtet. Vorerst werden nur *C* und *P* im 4-Tupel betrachtet und die möglichen Kombinationen aufgelistet (siehe Tabelle 1).

Tabelle. 1: Mögliche Kombinationen von *C* und *P* in einem 4-Tupel.

Tupel	Beschreibung [1]
(_,_,C,P)	Ersteller und Patient sind bekannt
(_,_,C,P)*	Ersteller ist bekannt und Patient ist pseudonymisiert
(_,_,C,_)	Ersteller ist bekannt und Patient ist anonymisiert
(_,_,_,P)	Ersteller ist nicht bekannt, aber Patient ist bekannt
(_,_,_,P)*	Ersteller ist nicht bekannt und Patient ist pseudonymisiert
(_,_,_,_)	Vollständige Anonymisierung

Durch die Variation von C und P kann sowohl der Ersteller und Patient transparent hinterlegt werden, als auch eine vollständige Anonymisierung dieser erreicht werden. Für die Datenfreigabe werden nun die Szenarien G2G, G2P, P2I und P2G diskutiert (siehe Tabelle 2).

Tabelle 2: Mögliche Szenarien für GDA , P und I.

	Sender	**Empfänger**	
	Speicherung	Übermittlung	Speicherung
G2G	**GDA**	**GDA**	
1	*(S,R,C,P)*	*(S,R,C,P)*	*(S,R,C,P)*
2	*(S,R,C,P)*	*(S,R,_,_)*	*(S,R,_,_)*
G2P	**GDA**	**P**	
3	*(S,R,C,P)*	*(S,R,C,P)*	*(S,R,C,P)*
P2G	**P**	**GDA**	
4	*(S,R,C,P)*	*(S,R,C,P)*	*(S,R,C,P)*
5	*(S,R,C,P)*	*(S,R,_,P)*	*(S,R,_,P)*
P2I	**P**	**I**	
6	*(S,R,C,P)*	*(S,R,C,P)*	*(S,R,C,P)*
7	*(_,_,_,_)*	*(_,R,C,P)*	*(_,R,C,P)*

Dabei wird auf Seite des Senders das 4-Tupel herangezogen, das beim Sender gespeichert wird. Beim Empfänger wird einerseits das 4-Tupel, das der Sender an den Empfänger übermittelt hat und andererseits das 4-Tupel, das der Empfänger hinterlegt hat, betrachtet. Die Fälle 1, 3, 4 und 6 in Tabelle 2 sind die Standardfälle, wobei jeweils die vollständige Information gespeichert wird. Im Fall 2 werden die vollständigen Informationen beim Sender-GDA hinterlegt, jedoch dem Empfänger-GDA in anonymisierter Form übermittelt (z.B. Einholen eines Zweitgutachtens (Second opinion) für einen anonymen Patienten). Im Fall 5 übermittelt ein Patient eine Referenz auf ein Dokument, wobei der Ersteller des Dokuments verborgen bleibt (z.B. Einholen einer Zweitmeinung). Der Fall 7 entspricht der Freigabe von Daten zwischen Teil-Identitäten. Dabei speichert die freigebende Teil-Identität keine Informationen, um die Nachvollziehbarkeit zu unterbinden und übermittelt der anderen Teil-Identität ein Tupel ohne die Senderinformationen. Dadurch ist die übermittelnde Partei auch beim Empfänger nicht eruierbar. Enthält das übermittelte 4-Tupel Informationen über den Ersteller *C*, so sollte dieser als Sender eingetragen werden. Dadurch wird das 4-Tupel vollständig und man unterbindet etwaige Fragen bezüglich „dem Fehlen des Senders“.

[1] Unter der Voraussetzung, dass im referenzierten Dokument keine ersteller- und patientenbezogenen Informationen enthalten sind.

Für die Verwaltung der Verzeichnisse der involvierten Parteien und der Freigaben werden im Allgemeinen Datenbanksysteme eingesetzt. Um im Gegensatz zu herkömmlichen Systemen die Verknüpfungen zwischen Datenobjekten (z.B. Patient, Freigabe) in der Datenbank zu unterbinden, können sogenannte pseudonymisierte Beziehungen eingesetzt werden [2,3]. Wird beispielsweise die Beziehung zwischen Patienten und Freigaben pseudonymisiert, so kann ein Patient genau nur seine Freigaben identifizieren und interpretieren. Dies wird durch die Verwendung kryptographischer Methoden realisiert. Für alle anderen Personen im System ist die Zuordnung zwischen dem Patienten (privaten Identitäten) und seinen Freigaben in keinster Weise nachvollziehbar. Dieser Zusammenhang wird in Abbildung 3 anhand eines einfachen, schematischen Beispiels erläutert.

ID_1	USERNAME	ID_2	ID_USER
1	$USER_1$	-	-
-	-	$E(r_2, PIN_2)$	1
-	-	$E(r_3, PIN_3)$	1
-	-	$E(r_4, PIN_4)$	1
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

PATIENT

1 n

ID	REF	4-TUPEL
1	$E(D_5, OS_1)$	$E((S,\_,C,P), OS_1)$
1	$E(D_6, OS_1)$	$E((S,\_,C,P), OS_1)$
r_2	$E(D_5, PIN_2)$	$E((\_,\_,C,P), PIN_2)$
r_3	$E(D_5, PIN_3)$	$E((\_,\_,C,P), PIN_3)$
r_3	$E(D_6, PIN_3)$	$E((\_,\_,C,P), PIN_3)$
⋮	⋮	⋮

FREIGABE

Abbildung 3: Patient und seine Freigaben.

In diesem Beispiel wird der Ausschnitt eines Patienten und seiner Identitäten aus dem Patientenverzeichnis (PATIENT) dargestellt. Die erste Zeile symbolisiert die öffentliche Identität des Patienten und wird durch die *ID_1* mit dem Wert 1 identifiziert. Jede private Identität (aktiv oder inaktiv) ist über den Identifikator der öffentlichen Identität mit dieser verknüpft. Weiters wird diese durch eine Zufallszahl r_i repräsentiert und mit der PIN der jeweiligen Identität verschlüsselt. Bei inaktiven Identitäten wird als PIN eine Zufallszahl herangezogen. Bei der Aktivierung einer inaktiven Identität wird dieser Repräsentant neu gewählt und mit einer vom Benutzer gewählten PIN verschlüsselt. In der zweiten Tabelle (FREIGABE) werden die Freigaben für diesen Patienten dargestellt. Dabei werden einerseits der Identifikator der öffentlichen Identität bzw. die Zufallszahlen der privaten Identitäten, eine verschlüsselte Referenz auf die Gesundheitsdaten (Dokument D_i) und die restlichen drei Komponenten des 4-Tupels dargestellt. Im Beispiel ist zu erkennen, dass eine Freigabe einer Identität ausschließlich dann dem Patienten zuordenbar ist, wenn die zur Entschlüsselung notwendige PIN bekannt ist. Für die kryptographischen Operationen wird eine Public-Key Infrastruktur (PKI) vorausgesetzt, damit die benötigten Schlüsselpaare zur Verfügung stehen. Die restlichen Informationen aus dem 4-Tupel werden für die öffentliche Identität mit dem öffentlichen Schlüssel (OS_i) des Patienten verschlüsselt, da grundsätzlich jede andere Partei eine Freigabe für den Patienten erstellen können soll. Bei den privaten Identitäten wird diese Information mittels der PIN dieser Identität symmetrisch verschlüsselt, da Formate für asymmetrische Verschlüsselung (z.B. PKCS#7 [4], XMLEnc [5]) Informationen über den Empfänger (in diesem Beispiel den Patienten) beinhalten können (siehe [2]). Es sei jedoch darauf hingewiesen, dass der zuvor beschriebene Ansatz lediglich zur Veranschaulichung einer pseudonymisierten Beziehung dient. Beispielsweise sollte eine Verschlüsselung einer Zufallszahl nicht ausschließlich mit einer PIN durchgeführt werden.

In diesem schematischen Ansatz könnte zudem aus der Kenntnis einer einzigen Freigabe einer privaten Identität auf alle Referenzen, die dieser zugeordnet sind, geschlossen werden. Außerdem sind generell statistische Auswertungen über Identifikatoren der Freigaben möglich. Diese Problematik kann jedoch durch die Anwendung einer Obfuscation-Technik (z.B. eines schwachen Zufallszahlengenerators) entschärft werden [6,7]. Dabei werden die Identifikatoren der privaten Identitäten

(Zufallszahlen r_i) nicht eindeutig gewählt, sodass Duplikate auftreten und folglich statistische Analysen kaum sinnvoll interpretierbare Ergebnisse liefern. Diese Technik bedingt einen Overhead an nicht erfolgreichen Entschlüsselungsoperationen. Durch geeignete Parametrierung der Obfuscation-Technik kann dieser jedoch systemweit auf ein vertretbares Maß beschränkt werden.

3. Ausblick

Zentrale eHealth-Portale werden das Gesundheitswesen durch einen orts- und zeitunabhängigen Zugriff auf Gesundheitsdaten nachhaltig verbessern. Nichtsdestotrotz müssen Aspekte wie beispielsweise die Kommunikationsinfrastruktur, die Interoperabilität und im Speziellen der Datenschutz betrachtet werden. Im Bereich des Datenschutzes wird im Allgemeinen eine Vielzahl von Angriffsszenarien berücksichtigt, jedoch wird die erzwungene Offenlegung kaum betrachtet. Dieser Aspekt wurde in der vorliegenden Arbeit eingehend beleuchtet und Lösungsansätze vorgestellt. Weiterführende Themen in diesem Bereich sind eine verbesserte Authentifizierungsmöglichkeit der Benutzer, sowie eine konkrete Implementierung in Form einer webbasierten browser-unabhängigen Softwarelösung.

4. Referenzen

[1] Machbarkeitsstudie ELGA, 2006, http://www.bmgf.gv.at

[2] C. Stingl, D. Slamanig, et al. Realisierung eines sicheren zentralen Datenrepositories. In (Patrick Horster, Ed.) DACH Security, pp. 32-45, IT Verlag, 2006.

[3] A. Pfitzmann, M. Köhntopp. Anonymity, Unobservability, and Pseudonymity - A Proposal for Terminology. Workshop on Design Issues in Anonymity and Unobservability, LNCS, vol. 2009, Springer, 2000, pp. 1-9.

[4] PKCS#7, Cryptographic Message Syntax Standard, PKCS, RSA Laboratories, 1993.

[5] XML Encryption Syntax and Processing. W3C Recommendation 10 December 2002 (http:// www.w3.org/TR/2002/REC-xmlenc-core-20021210/)

[6] D.E. Bakken, R. Parameswaran, D.M. Blough, A.A. Franz, and T.J. Palmer. Data Obfuscation: Anonymity and Desensitization of Usable Data Sets. IEEE Security and Privacy 2, 6 (Nov. 2004), pp. 34-41.

[7] S.D. Damiani, S. Vimercati, S., et al.: Balancing confidentiality and efficiency in untrusted relational DBMSs. In Proceedings of the 10th ACM Conference on Computer and Communications Security CCS '03, pp. 93-102.