



Institut für Qualitätssicherung und
Transparenz im Gesundheitswesen

Systemweit einheitliche Leistungserbringer- pseudonymisierung

Anwenderhandbuch Pseudonymisierungsprogramm V02

Erstellt im Auftrag des
Gemeinsamen Bundesausschusses

Stand: 09. Februar 2017

Impressum

Thema:

Pseudonymisierungsprogramm zur systemweit einheitlichen Leistungserbringerpseudonymisierung. Anwenderhandbuch V02

Auftraggeber:

Gemeinsamer Bundesausschuss

Datum der Abgabe:

09. Februar 2017

Herausgeber:

IQTIG – Institut für Qualitätssicherung
und Transparenz im Gesundheitswesen

Katharina-Heinroth-Ufer 1
10787 Berlin

Telefon: (030) 58 58 26-0
Telefax: (030) 58 58 26-999

info@iqtig.org

<http://www.iqtig.org>

Inhaltsverzeichnis

Tabellenverzeichnis	4
Abbildungsverzeichnis	4
Informationen zu diesem Dokument.....	5
1 Einleitung.....	6
1.1 Verschlüsselungsverfahren zur Pseudonymisierung.....	6
1.2 Anwendungsbereich.....	7
1.3 Eigenschaften der LE-Pseudonymisierung.....	8
2 Installation.....	9
2.1 Systemvoraussetzungen.....	9
2.2 Programmbestandteile.....	9
2.3 Referenzimplementierung.....	9
2.4 Installation durchführen.....	9
2.5 Anpassung der Konfigurationsdatei.....	10
2.6 Programm-Updates.....	10
3 Programmaufbau	11
3.1 Grafische Benutzeroberfläche.....	11
3.1.1 Vorkonfiguration des Programms.....	11
3.1.2 Schlüsselpaargenerierung.....	13
3.1.3 Durchführung der Pseudonymisierung.....	15
3.1.4 Durchführung der Depseudonymisierung.....	15
3.1.5 Übertragung der Pseudonyme an die BAS.....	16
3.2 Konsolenanwendung.....	17
3.3 Zusatzdatenverwaltung.....	19
4 Fehlermeldungen	23
Anhang.....	27

Tabellenverzeichnis

Tabelle 1: Eingabe- und Auswahlfelder der Pseudonymisierungs-, Depseudonymisierungs- und Konfigurationsoberfläche	11
Tabelle 2: Eingabe- und Auswahlfelder der Oberfläche zur Schlüsselpaarerzeugung.....	13
Tabelle 5: Eingabe- und Auswahlfelder der Oberfläche zur Übertragung der Zusatzinformationen an die BAS	16
Tabelle 6: Anwendungsfälle der Konsolenanwendung.....	17
Tabelle 7: Parameter der Konsolenanwendung.....	18
Tabelle 8: Spaltenbedeutung der CSV-Datei	20
Tabelle 9: Fehlercodes	23
Tabelle 11: Gültige Länderkürzel.....	27
Tabelle 12: Gültige Datenannahmestellen.....	27

Abbildungsverzeichnis

Abbildung 1: Konfigurationsoberfläche des LE-Pseudonymisierungsprogrammes.....	13
Abbildung 2: Oberfläche zur Generierung der Schlüsselpaare.....	14
Abbildung 3: Oberfläche zur Pseudonymisierung der LE-Informationen.....	15
Abbildung 4: Oberfläche zur Depseudonymisierung der LE-Informationen.....	16
Abbildung 5: Oberfläche zur Übermittlung der Zusatzinformationen an die BAS.....	17

Informationen zu diesem Dokument

Darstellungsmittel

Im Folgenden sind Symbole und Darstellung besonderer Informationen beschrieben.



Achtung

Beschreibt Ursache, Folge und Vermeidung einer besonderen Fehlanwendung, die zu Problemen bei der Implementierung oder Ähnlichem führen kann.

Beispiel:

Beispiele sind ein Hilfsmittel, um zuvor vermittelte Informationen oder konkrete Abschnitte der Anwendung zu verdeutlichen.

Zielgruppe

Dieses Handbuch richtet sich an administrative Mitarbeiter der Datenannahmestellen (DAS) und Softwareentwickler, die mit der Umsetzung der Leistungserbringerpseudonymisierung beschäftigt sind.

Änderungen in der Version 02 (gegenüber Version 2015 V1.0)

- Anpassung des Dokuments an das Layout des Instituts nach § 137a SGB V (IQTIG)
- Redaktionelle Änderungen

Änderungen in der Version 03 (gegenüber Version 02)

- Anpassung der Systemvoraussetzungen auf Oracle Java 8

Änderungen in der Version 04 (gegenüber Version 03)

- Anpassung der Tabelle mit den gültigen Datenannahmestellen

1 Einleitung

Die durch die Richtlinie zur einrichtungs- und sektorenübergreifenden Qualitätssicherung (Qesü-RL)¹ vorgenommene Ausdehnung der Qualitätssicherung vom stationären auf den vertragsärztlichen Bereich sowie die Einbeziehung von Sozialdaten bei den Krankenkassen in die Qualitätsauswertungen erfordern eine einheitliche Pseudonymisierung der leistungserbringeridentifizierenden Daten. Die systemweit einheitliche Pseudonymisierung bildet die Grundlage für die Zusammenführung der Qualitätssicherungsdaten (QS-Daten) aus den unterschiedlichen Datenquellen in der Bundesauswertungsstelle (BAS) zur Berechnung der Qualitätsindikatoren sowie für den Versand der Rückmeldeberichte von der BAS an die Leistungserbringer (LE). Die Qesü-RL fordert die systemweit einheitliche Leistungserbringerpseudonymisierung (LE-Pseudonymisierung) in § 3 Abs. 2 Satz 4 der Anlage zu Teil 1 der Richtlinie:

Für die Pseudonymisierung stimmen die Datenannahmestellen nach § 9 Absatz 1 Satz 2 der Richtlinie (KV² bzw. KZV³), die Datenannahmestellen nach § 9 Absatz 1 Satz 3 der Richtlinie (LQS⁴/LKG⁵) sowie die Datenannahmestelle nach § 9 Absatz 1 Satz 5 der Richtlinie (DAS-KK⁶) untereinander ein Verfahren ab, welches sicherstellt, dass die Datenannahmestellen den gleichen leistungserbringeridentifizierenden Daten jeweils das gleiche Pseudonym zuordnen.

Die Datenflussspezifikation zur systemweit einheitlichen Leistungserbringerpseudonymisierung beschreibt die Datenflüsse, Dateiformate zum Datenaustausch sowie algorithmische Grundlagen der LE-Pseudonymisierung. Das Dokument richtet sich an die o. g. Datenannahmestellen und an die mit der Umsetzung beauftragten Softwarehersteller.

Die Erzeugung von kryptografischen Schlüsseln und von Zertifikaten und die Pseudonymisierung sowie Depseudonymisierung von leistungserbringeridentifizierenden Daten werden mithilfe des Leistungserbringerpseudonymisierungsprogramms durchgeführt. Die Anwenderdokumentation für das Programm befindet sich in der Datei „Pseudonymisierungsprogramm.pdf“ und ergänzt die Datenflussspezifikation. Auf der Website <http://www.iqtig.org> kann die aktuelle Version des Programms, der Anwenderdokumentation und der Datenflussspezifikation als ZIP-Archiv heruntergeladen werden.

1.1 Verschlüsselungsverfahren zur Pseudonymisierung

Die Erstellung der LE-Pseudonyme wird mittels eines asymmetrischen (Public-Key-)Verschlüsselungsverfahrens durchgeführt. Im Gegensatz zu symmetrischen Verschlüsselungsverfahren wird kein gemeinsamer geheimer Schlüssel benötigt, den beide Kommunikationspartner kennen. Bei

¹ <https://www.g-ba.de/informationen/richtlinien/72/>

² Kassenärztliche Vereinigung (KV)

³ Kassenzahnärztliche Vereinigung (KZV)

⁴ Landesgeschäftsstelle für Qualitätssicherung (LQS)

⁵ Landeskrankenhausgesellschaft (LKG)

⁶ Datenannahmestelle für die Krankenkassen (DAS-KK)

der asymmetrischen Verschlüsselung wird vom Benutzer (LQS/LKG/KV/KZV) ein Schlüsselpaar, bestehend aus einem privaten Schlüssel und einem öffentlichen Schlüssel, generiert. Die Schlüssel sind gemeinsam mit weiteren Informationen in Zertifikate eingebettet. Der private Schlüssel wird bei der Erzeugung des Schlüsselpaars mit einem Passwort geschützt. Der öffentliche Schlüssel wird an alle Kommunikationspartner verteilt. Mithilfe des öffentlichen Schlüssels können Daten verschlüsselt und sicher an den Besitzer des privaten Schlüssels übermittelt werden. Mit dem privaten Schlüssel und dem Passwort können die empfangenen Daten entschlüsselt werden. Die öffentlichen Schlüssel sind in Zertifikate eingebettet, die zusätzliche Metadaten enthalten.

Für die LE-Pseudonymisierung wird ein deterministisches, asymmetrisches Verschlüsselungsverfahren gebraucht. Verwendung findet der RSA-Verschlüsselungsalgorithmus von *Legion of the Bouncy Castle Inc*⁷. Dieser Algorithmus erfüllt sowohl die Anforderungen an die systemweit einheitliche Pseudonymisierung als auch die vom Bundesamt für Sicherheit in der Informationstechnik (BSI) empfohlenen Kriterien. Die deterministische Verschlüsselung nimmt in diesem Zusammenhang einen besonderen Stellenwert ein, da sichergestellt ist, dass die Verschlüsselung derselben LE-identifizierenden Daten mit dem gleichen öffentlichen Schlüssel bei verschiedenen DAS immer das gleiche Ergebnis produziert.



Achtung

Die Zertifikate mit den öffentlichen Schlüsseln dürfen nur den im Datenfluss vorgesehenen Stellen zur Verfügung gestellt werden. Es ist z. B. falsch, diese Zertifikate auf einer Website oder mit der Spezifikation zu veröffentlichen. In diesem Fall würde durch die Nutzung der deterministischen Verschlüsselung ein Besitzer von IKNR/BSNR die jeweiligen Pseudonyme erstellen können. Das Zertifikat, welches den privaten Schlüssel enthält, dient nur dazu, Rückmeldeberichte mit Sozialdaten dem richtigen Leistungserbringer zuordnen zu können, und darf daher Dritten nicht zur Verfügung gestellt werden.

1.2 Anwendungsbereich

Das Pseudonymisierungsprogramm wird für unterschiedliche Zielgruppen und Anwendungsmöglichkeiten in folgenden Varianten bereitgestellt:

- mit einer grafischen Benutzeroberfläche (siehe Abschnitt 1.1)
- als Konsolenanwendung (siehe Abschnitt 1.2)
- als Programmierschnittstelle (siehe Abschnitt 1.3)

Die grafische Benutzeroberfläche wurde für Testzwecke und die manuelle Integration der Leistungserbringerpseudonymisierung in den Datenfluss einer Datenannahmestelle erstellt. Eine Form der manuellen Integration wäre z. B., wenn eine DAS alle bekannten IKNR/BSNR mithilfe der Benutzeroberfläche pseudonymisiert, eine Mappingtabelle zwischen Pseudonym und Klar-text erstellt und diese dann zu Ver- und Entschlüsselung einsetzt.

⁷ <https://www.bouncycastle.org/>

Die Konsolenanwendung wurde für die programmatische Integration der Leistungserbringerpseudonymisierung in nicht Java-basierte Datenflüsse bei den Datenannahmestellen erstellt.

Zusätzlich steht den Java-Entwicklern bei den Datenannahmestellen eine API (Application Programming Interface/Programmierschnittstelle) zur Verfügung.

Zusatzinformationen zum Leistungserbringer wie Standorte oder Fachabteilungen können mit dem Tool über eine Webservice-Schnittstelle (SOAP) in Form von CSV-Dateien an die Bundesauswertungsstelle übertragen werden (siehe Abschnitt 1.1).

1.3 Eigenschaften der LE-Pseudonymisierung

Das Pseudonymisierungsprogramm nutzt einen deterministischen, asymmetrischen RSA-Verschlüsselungsalgorithmus mit einer Schlüssellänge von 2048 Bit zur Erzeugung von Pseudonymen. Die resultierenden Binärdaten werden dann in einen Base64-kodierten Text überführt und als Pseudonym benutzt.

Ein LE-Pseudonym ist folgendermaßen aufgebaut:

- **Art der Datenannahme** (BAS/LQS/KV/SV)
- **Bundesland/Region**
- spezifizierte Trennzeichen (\$\$, ##)
- **Mit dem öffentlichen Schlüssel verschlüsselte IKNR/BSNR**

Beispiel:

LQS\$\$HH##fL1PqnoBIIAaM2bXSXvf1MOnhp4NhPx2sKqnIng

Die einheitliche Pseudonymisierung ist von allen Datenannahmestellen gemäß Qesü-RL durchzuführen. Für die Krankenhäuser ist die Institutskennzeichen-Nummer (IKNR) zu pseudonymisieren, im kollektiv- bzw. selektivvertraglichen Bereich ist stattdessen die Betriebsstättennummer (BSNR) (XML-Element `BSNRAMBU`) zu pseudonymisieren. Die DAS für die Krankenkassen und die DAS für die selektivvertraglichen Arztpraxen (DAS-SV) besitzen keine eigenen Zertifikate und müssen deshalb die öffentlichen Zertifikate der DAS-LE auf Landesebene nutzen. Dazu müssen diese DAS entscheiden, ob es sich bei den leistungserbringeridentifizierenden Daten um eine IKNR oder BSNR handelt. Die IKNRs werden mit dem Schlüssel der LQS verschlüsselt, die BSNRs mit dem Schlüssel der KV. Zusätzlich muss auch das Bundesland bekannt sein, damit der richtige Schlüssel auf Landesebene ausgewählt werden kann.

Während der Durchführung der Pseudonymisierung wird das QS-Verfahren dem Pseudonymisierungsprogramm als Parameter übergeben. Dies geschieht beispielsweise in der CSV-Datei, die auf der grafischen Oberfläche als Eingabedatei angegeben werden kann. Das Pseudonymisierungsprogramm enthält in der jeweils aktuellsten Version alle gültigen QS-Verfahren, bei denen die systemweit einheitliche LE-Pseudonymisierung angewendet wird und prüft die Eingabedaten auf ihre Korrektheit. Somit wird verhindert, dass durch fehlerhafte Angaben des QS-Verfahrens in einer DAS ungültige LE-Pseudonyme erstellt werden.

2 Installation

2.1 Systemvoraussetzungen

Hardware

- mindestens 128 MB Arbeitsspeicher (RAM)
- mindestens 25 MB freier Festplattenspeicher

Software

- Betriebssystem, welches die Installation und Ausführung von Oracle Java ermöglicht
- Oracle Java 8
(siehe auch <https://www.java.com/de/download/help/sysreq.xml>)

2.2 Programmbestandteile

Das Pseudonymisierungsprogramm mit der grafischen Benutzeroberfläche besteht aus zwei Dateien:

- Ausführbares Java-Programm
- Eigenschaftsdatei im Benutzer-Heim-Verzeichnis

Weitere Dateien sind für das Pseudonymisierungsprogramm bei der Datensicherung oder Datenreorganisation nicht zu berücksichtigen.

2.3 Referenzimplementierung

Zur Vereinfachung der Einbindung der API in Softwareentwicklungsprojekte wurde diese mit ihrer Referenzimplementierung als Maven-Artefakt bereitgestellt, sodass sie in das entsprechende lokal verwendete Repository deployed werden kann. Die Referenzimplementierung des Pseudonymisierungsprogrammes befindet sich in `pseudonymisierung-impl.jar`. Die Referenzimplementierung der Zusatzdatenverwaltung ist in `pseudonymisierung-zus-cli.jar` zu finden.

Sollte eine Datenannahmestelle die API verwenden wollen, ohne Maven einzusetzen, so finden sich die Informationen über abhängige Bibliotheken in den `pom.xml`-Dateien.

Alle benötigten Dateien befinden sich im Verzeichnis

`Programm\API_und_Referenzimplementierung`
innerhalb des PSP-ZIP-Archivs.

2.4 Installation durchführen

Java

Das Pseudonymisierungsprogramm wurde in der Programmiersprache Java entwickelt. Zur Ausführung von Java-Programmen wird das Java Runtime Environment (JRE) zur Verfügung gestellt. Das JRE kann „online“ und „offline“ installiert werden.

Offizielle Installationspakete können von der Website <http://www.java.com/de/download/manual.jsp> heruntergeladen werden.

Anleitungen zur Installation von Java unter verschiedenen Betriebssystemen sind auf der folgenden Website detailliert beschrieben:

https://www.java.com/de/download/help/download_options.xml

Pseudonymisierungsprogramm

Die grafische Benutzeroberfläche für das Pseudonymisierungsprogramm ist ein ausführbares Java-Programm und kann bei korrekt installierter Java-Laufzeitumgebung einfach mit einem Doppelklick unter Windows-Systemen gestartet werden.

Das Konsolenprogramm kann über die Windows-Kommandozeile aufgerufen werden.

2.5 Anpassung der Konfigurationsdatei

Die grafische Benutzeroberfläche speichert die Konfiguration des Programms in einer Eigenschaftsdatei (Properties). Diese wird im Benutzer-Heim-Verzeichnis angelegt.

Zum Beispiel unter Windows 7:

`C:\Users\M.Mustermann\pseudonym-config.properties`

Beim ersten Start des Programms wird diese Datei aus den mitgelieferten Initialeinstellungen befüllt und kann per Texteditor oder auf der Konfigurationsseite der grafischen Benutzeroberfläche angepasst werden.

2.6 Programm-Updates

Das Pseudonymisierungsprogramm wird in unregelmäßigen Intervallen Updates erhalten. Gründe für Updates können Fehlerbehebungen, Programmverbesserungen sowie die Anpassung des Programms an neue oder geänderte QS-Verfahren sein. Die alte Programmversion kann durch die neue Version ersetzt werden, indem die JAR-Datei ausgetauscht wird. Es besteht auch die Möglichkeit, mehrere Programmversionen parallel zu nutzen. Jede Programmversion nutzt die gemeinsame Konfigurationsdatei im Benutzer-Heim-Verzeichnis, sodass einmal vorgenommene Einstellungen nicht verloren gehen oder erneut durchgeführt werden müssen.

3 Programmaufbau

Das Pseudonymisierungsprogramm wird in unterschiedlichen Varianten bereitgestellt, deren Funktionsweise in den nächsten Unterkapiteln genauer erläutert wird.

3.1 Grafische Benutzeroberfläche

Datenannahmestellen können die grafische, interaktive Benutzeroberfläche nutzen, um das Pseudonymisierungsprogramm zu konfigurieren und Pseudonymisierungs- bzw. Depseudonymisierungsprozesse zu starten. Mit dem Programm können zudem weitere Zusatzinformationen an die Bundesauswertungsstelle (BAS) übermittelt werden.

3.1.1 Vorkonfiguration des Programms

Bei der erstmaligen Verwendung des Pseudonymisierungsprogramms ist es zur Verringerung redundanter Datenübergaben sinnvoll, die wiederkehrenden Dateneingaben vorab zu konfigurieren. Eine Überschreibung der Standardkonfiguration ist bei jedem Anwendungsfall möglich.

Beim ersten Start wird die Konfigurationsdatei `pseudonym-config.properties` im Benutzer-Heim-Verzeichnis abgelegt. Die hierbei angegebenen Spalten und Zeilennummern beginnen mit Spalte 0 und Zeile 1.

Tabelle 1: Eingabe- und Auswahlfelder der Pseudonymisierungs-, Depseudonymisierungs- und Konfigurationsoberfläche

Feld	Beschreibung
CSV-Datei-Spalte mit Eingangswert	<p>Pseudonymisierung: In dieser Spalte steht der zu pseudonymisierende Klartext (IKNR/BSNR).</p> <p>Depseudonymisierung: In dieser Spalte steht das Pseudonym.</p>
CSV-Datei-Spalte mit Ergebniswert	<p>Die Spalte der CSV-Datei, in die der Ergebniswert des Pseudonymisierungsvorgangs eingetragen wird.</p> <p>Ist der Wert größer als die Anzahl der Spalten in der CSV-Datei, wird das Ergebnis als letzte Spalte in die CSV-Datei eingetragen.</p> <p>Ist der Wert kleiner als 0, wird der Eingangswert mit dem Ergebnis überschrieben.</p>
CSV-Datei-Spalte mit Pseudonymisierungsverfahren	<p>Gibt an, in welcher Spalte der CSV-Datei die Pseudonymisierungsmethode angegeben ist. Wird dieses Feld auf einen Wert kleiner 0 gestellt, so wird der Wert des Felds „Pseudonymisierungsverfahren“ für die gesamte CSV-Datei angewendet.</p>
Passwort des Benutzers bei der BAS	<p>Das Passwort, das bei der BAS dem entsprechenden Benutzernamen zugeordnet ist.</p>

Feld	Beschreibung
Datenannahmestellentyp	Typ der Datenannahmestelle auswählen (LQS, KV, BAS)
Benutzername bei der BAS	Jeder DAS, LQS, KV, welche Zusatzdaten verwaltet, wird von der BAS ein Benutzername zugeordnet.
Eingangs-CSV-Dateiname	Die CSV-Datei, welche die Eingangsdaten enthält.
Datum Gültigkeitsende des Schlüssels	Gibt an, bis zu welchem Datum das nächste erzeugte Schlüsselpaar gültig ist. Bleibt dieses Feld leer, ist das Schlüsselpaar ab der Erzeugung 20 Jahre lang gültig.
Ergebnis-CSV-Dateiname	Die CSV-Datei, in welcher das Pseudonymisierungsprogramm das pseudonymisierte bzw. depseudonymisierte Ergebnis ablegen soll.
Landeskennung	Auswahl der Landeskennung.
Passwort, mit dem der private Schlüssel geschützt wird	Der private Schlüssel wird verschlüsselt in einem X.509-Zertifikat abgelegt. Das Passwort wird benötigt, um den privaten Schlüssel benutzen zu können.
QS-Verfahren	Das QS-Verfahren, auf welches sich die Pseudonymisierung bezieht. Der Wert des Dropdown-Menüs wird herangezogen, wenn bei „CSV-Datei-Spalte mit Pseudonymisierungsverfahren“ der Wert kleiner 0 ist.
Schlüsselverzeichnis	Das Verzeichnis, in dem die kryptografischen Schlüssel abgelegt sind.
Datum Gültigkeitsstart des Schlüssels	Gibt an, ab welchem Datum das nächste erzeugte Schlüsselpaar gültig ist.
Zeilennummer der CSV-Datei, mit der die Pseudonymisierung beginnen soll	Für den Fall, dass nicht die gesamte CSV-Datei pseudonymisiert werden soll, oder wenn es eine Kopfzeile mit Spaltennamen in der CSV-Datei gibt, kann angegeben werden, ab welcher Zeile die CSV-Datei bearbeitet werden soll.
Zusatzdaten CSV-Dateiname	Die CSV-Datei mit den Zusatzinformationen (Standortnummern, Haupt-IKNR/BSNR, Gültigkeit)

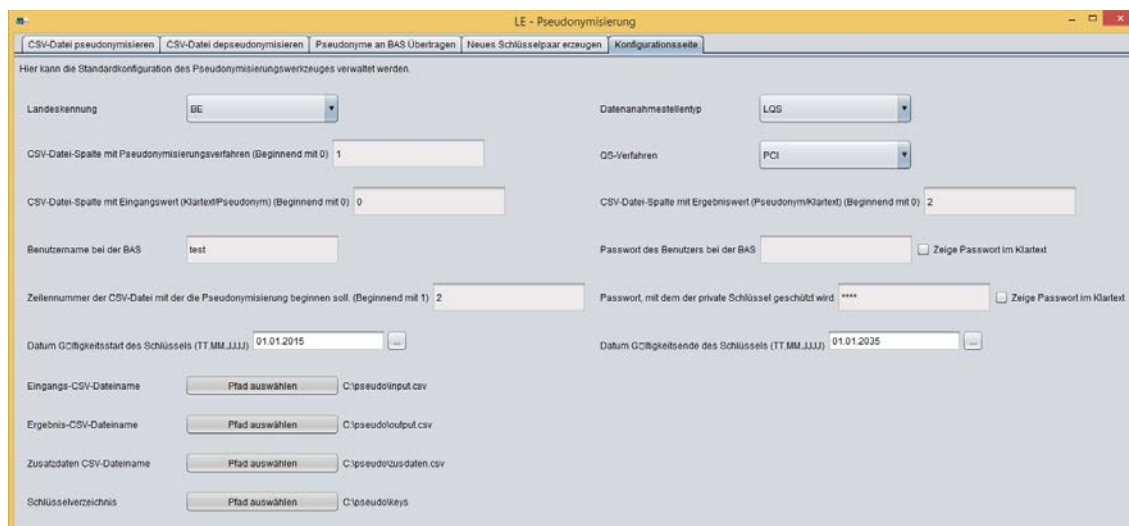


Abbildung 1: Konfigurationsoberfläche des LE-Pseudonymisierungsprogrammes

3.1.2 Schlüsselpaargenerierung

Bei der Generierung der Schlüsselpaare (X.509-Zertifikate) sind folgende Hinweise zu berücksichtigen:

- Nur die Datenannahmestellen der Krankenhäuser (LQS/LKG) und die Kassenärztlichen Vereinigungen (KV) bzw. Kassenzahnärztlichen Vereinigungen (KZV) dürfen öffentliche und private Zertifikate generieren
- Die DAS-KK und DAS-SV erhält für den Regelbetrieb von den LQS/LKG/KV/KZV über einen sicheren Kanal die öffentlichen Zertifikate
- Private Zertifikate dürfen nur vom Zertifikatsbesitzer selbst für die Depseudonymisierung verwendet werden
- Die Datenflüsse zur Übermittlung der Zertifikate zwischen den einzelnen Stellen werden in der Datenflussspezifikation beschrieben

Während der Generierung der Zertifikate werden verschiedene Attribute abgefragt (siehe Tabelle 2).

Tabelle 2: Eingabe- und Auswahlfelder der Oberfläche zur Schlüsselpaarerzeugung

Feld	Beschreibung
Datenannahmestellentyp	Art der Datenannahmestelle auswählen (LQS, KV, BAS).
Landeskennung	Auswahl der Landeskennung.
Startgültigkeitsdatum des Schlüssels	Gibt an, ab welchem Datum das nächste erzeugte Schlüsselpaar gültig ist.
Endgültigkeitsdatum des Schlüssels	Gibt an, bis zu welchem Datum das nächste erzeugte Schlüsselpaar gültig ist. Bleibt dieses Feld leer, so bleibt das Schlüsselpaar ab der Erzeugung 20 Jahre lang gültig.

Feld	Beschreibung
	Wir empfehlen, den Gültigkeitszeitraum nicht zu kurz zu wählen, da sich mit jedem neuen Schlüssel die Pseudonyme ändern und erneut Listen von zusammengehörigen Pseudonymen an die BAS übertragen werden müssen.
Passwort, mit dem der private Schlüssel geschützt wird	Der private Schlüssel wird verschlüsselt in einem X.509-Zertifikat abgelegt. Das Passwort wird benutzt, um den privaten Schlüssel benutzen zu können. Sollte dieses Passwort verloren gehen, müssen die Daten des laufenden Erfassungsjahres repseudonymisiert und erneut an die BAS übertragen werden.
Schlüssel Verzeichnis	Das Verzeichnis, in dem die Schlüssel abgelegt werden.

Die Ausgabe besteht aus zwei Dateien, dem öffentlichen und dem privaten Zertifikat. Die Dateien werden automatisch nach folgendem Schema benannt:

▪ Das öffentliche Zertifikat:

<Art-Datenannahme>\$\$<Bundeslandkürzel><Startgültigkeitsdatum>.cer

▪ Das private Zertifikat:

<Art-Datenannahme>\$\$<Bundeslandkürzel><Startgültigkeitsdatum>.p12

Beispiel (LQS Hamburg):

Öffentliches Zertifikat: LQS\$\$HH20150501.cer

Privates Zertifikat: LQS\$\$HH20150501.p12

Abbildung 2: Oberfläche zur Generierung der Schlüsselpaare

Beim Klick auf die Schaltfläche „Zurücksetzen“ werden die Formularwerte auf die Werte der Konfigurationsseite zurückgesetzt.

3.1.3 Durchführung der Pseudonymisierung

Das Pseudonymisierungsprogramm nutzt den in Abschnitt 1.1 beschriebenen Algorithmus zur Durchführung der LE-Pseudonymisierung. Die Oberfläche zur Pseudonymisierung ist in Abbildung 3 dargestellt. Eine detaillierte Beschreibung der entsprechenden Auswahlfelder befindet sich in Tabelle 1.

Weitere Informationen über die LE-Daten und den Aufbau der Exportdateien sind jeweils der technischen Dokumentation der Basisspezifikation für Leistungserbringer bzw. Datenannahmestellen zu entnehmen.

Abbildung 3: Oberfläche zur Pseudonymisierung der LE-Informationen

Beim Klick auf die Schaltfläche „Zurücksetzen“ werden die Formularwerte auf die Werte der Konfigurationsseite zurückgesetzt.

3.1.4 Durchführung der Depseudonymisierung

Analog zur Pseudonymisierung nutzt das Pseudonymisierungsprogramm denselben Algorithmus zur Depseudonymisierung der Leistungserbringerpseudonyme. Die Depseudonymisierung mit dem Pseudonymisierungsprogramm ist nur durch die LQS/LKG/KV/KZV durchzuführen. Die DAS-KK und DAS-SV können hingegen nur auf Mapping-Informationen zurückgreifen. Die Depseudonymisierung der LE-Pseudonyme in den Rückprotokollen bzw. in den Rückmeldeberichten ist gemäß Richtlinie erforderlich.

Die Oberfläche zur Depseudonymisierung ist in Abbildung 4 dargestellt. Eine detaillierte Beschreibung der entsprechenden Auswahlfelder befindet sich in Tabelle 1.

Abbildung 4: Oberfläche zur Depseudonymisierung der LE-Informationen

Beim Klick auf die Schaltfläche „Zurücksetzen“ werden die Formularwerte auf die Werte der Konfigurationsseite zurückgesetzt.

3.1.5 Übertragung der Pseudonyme an die BAS

Die Zusatzdaten, welche an die BAS übermittelt werden, sind ausführlich in Abschnitt 1.1 beschrieben.

Vor der Übermittlung der Daten an die BAS werden verschiedene Attribute abgefragt, die in Tabelle 5 näher erläutert werden.

Tabelle 3: Eingabe- und Auswahlfelder der Oberfläche zur Übertragung der Zusatzinformationen an die BAS

Feld	Beschreibung
Datenannahmestellentyp	Art der Datenannahmestelle auswählen (LQS, KV, BAS)
Landeskennung	Auswahl der Landeskennung.
Eingangs-CSV-Dateiname	Die zu bearbeitende CSV-Datei.
Zusatzdaten CSV-Dateiname	Die CSV-Datei mit den Zusatzinformationen (Standortnummern, Haupt-IKNR/BSNR, Gültigkeit)
Schlüssel Verzeichnis	Das Verzeichnis, in dem die Schlüssel abgelegt werden.
Der Benutzername bei der BAS	Jeder DAS, LQS, KV, welche Zusatzdaten verwaltet, erhält einen Benutzernamen von der BAS zugeordnet.
Passwort, mit dem der private Schlüssel geschützt wird	Der private Schlüssel wird verschlüsselt in einem X.509-Zertifikat abgelegt. Das Passwort wird benutzt, um den privaten Schlüssel benutzen zu können.

Abbildung 5: Oberfläche zur Übermittlung der Zusatzinformationen an die BAS

Beim Klick auf die Schaltfläche „Zurücksetzen“ werden die Formularwerte auf die Werte der Konfigurationsseite zurückgesetzt.

3.2 Konsolenanwendung

Die Datenannahmestellen können das Pseudonymisierungsprogramm als Konsolenapplikation in eine „Nicht-Java“-Datenverarbeitungskette integrieren. Das Programm wird in diesem Fall dafür benutzt, eine CSV-Datei mit einer Liste von LE-Daten zu pseudonymisieren bzw. zu depseudonymisieren. Hierbei wird ein System-Return-Code erzeugt, der die jeweiligen Erfolgs- oder Fehlerfälle ausgibt. Durch Konsolen-Parameter oder wahlweise eine Eigenschaftsdatei (Properties-File) erhält das Pseudonymisierungsprogramm sowohl die Information über den Zugriffsort der CSV-Datei als auch über den Speicherort der Ergebnisdatei.

Abhängig von den durchzuführenden Aktionen ist die Kommandozeile entsprechend anzupassen. Tabelle 6 gibt einen Überblick über die einzelnen Anwendungsfälle.

Tabelle 4: Anwendungsfälle der Konsolenanwendung

Use-Case	Kommandozeile
Generierung der Zertifikate	<code>java -jar pseudonymisierung-console.jar -a LQS -b 30.06.2015 -c C:\pseudo\key -l BW -s test -z</code>
Pseudonymisierung	<code>java -jar pseudonymisierung-console.jar -a LQS -c C:\pseudo\key -l NI -i C:\pseudo\input.csv -j 2 -k 1 -p -r C:\pseudo\out\outk.csv -t "-1" -w PCI</code>
Depseudonymisierung	<code>java -jar pseudonymisierung-console.jar -a LQS -c C:\pseudo\key -d -l NI -i C:\pseudo\depseu.csv -j 2 -k 1 -r C:\pseudo\out\outdepseuk.csv -t "-1" -s test</code>
Übermittlung der Zusatzdaten an die BAS	<code>java -jar pseudonymisierung-console.jar -h -a LQS -c C:\pseudo\key -l NI -I C:\pseudo\pseudonymMappings.csv -m C:\pseudo\zusatzDaten.csv -n username -q password</code>

Tabelle 5: Parameter der Konsolenanwendung

Parameter	Bedeutung
-a, --das <arg>	Datenannahmestelle, deren Zertifikate benutzt werden sollen. Gültige Werte sind LQS, BAS, KV, SV (siehe Anhang).
-b, --sdate <arg>	Beginn der Gültigkeit des zu erzeugenden Schlüssel-Zertifikats.
-c, --certificates <arg>	Verzeichnis für die Schlüsseldateien.
-d, --depseudo	Startet das Programm im Depseudonymisierungs-Modus.
-e, --edate <arg>	Ende der Gültigkeit des zu erzeugenden Schlüssel-Zertifikats.
-f, --cert <arg>	Zertifikat, welches verwendet werden soll.
-i, --input <arg>	CSV-Datei, welche pseudonymisiert bzw. depseudonymisiert werden soll. CSV-Datei, welche pro Zeile eine Liste von zusammengehörenden IKNR/BSNR enthält, für den Fall, dass Zusatzdaten an die BAS übertragen werden sollen.
-j, --rowStart <arg>	Zeile der CSV-Datei, ab der die Datei bearbeitet werden soll (Beginnend mit 1).
-k, --inputSpalte <arg>	Spalte in der CSV-Datei mit den Eingangsdaten (Klartext bei Pseudonymisierung, Pseudonym bei Depseudonymisierung), beginnend mit 0.
-l, --Länderkürzel <arg>	Länderkürzel (DAS-LE-Identität); siehe Tabelle 11
-o, --properties <arg>	Dateipfad der zu benutzenden Properties-Datei.
-p, --pseudo	Startet das Programm im Pseudonymisierungsmodus.
-r, --output <arg>	Ziel der Ergebnis-CSV-Datei
-s, --certpassword <arg>	Passwort der Zertifikatsdatei
-t, --outputSpalte <arg>	Die Spalte der CSV-Datei, in welche der Ergebniswert des Pseudonymisierungsvorganges eingetragen wird. Ist der Wert größer als die Anzahl der Spalten in der CSV-Datei, so wird das Ergebnis als letzte Spalte in die CSV-Datei eingetragen. Ist der Wert kleiner als 0, so wird der Eingangswert mit dem Ergebnis überschrieben.
-v, --verfahrenSpalte <arg>	Spalte in der CSV-Datei mit Verfahrenskürzel
-w, --verfahren <arg>	Verfahrenskürzel, welches benutzt werden soll.

Parameter	Bedeutung
-z, --creaCert	Erzeugt ein neues Zertifikat.
-h, --send	Sendet die aktuellen Einträge der Zusatzdatenverwaltung an die BAS.
-m, --addData <arg>	Die Zusatzdaten-CSV-Datei.
-n, --user <arg>	Der Benutzername bei der BAS.
-q, --password <arg>	Das Benutzerpasswort bei der BAS.

3.3 Zusatzdatenverwaltung



Achtung

Start der Übermittlung der Zusatzdaten voraussichtlich ab Oktober 2016

Zur Übermittlung der Zusatzdaten von der DAS an das IQTIG wird bei der DAS das Pseudonymisierungsprogramm benötigt, welches sich zu einem Webservice beim IQTIG verbindet. Der Webservice beim IQTIG wird voraussichtlich im **Oktober 2016** mit der Veröffentlichung des DIMDI-Updates der Basisspezifikation in Betrieb genommen werden. Bis zum genannten Zeitpunkt ist eine Übertragung der Zusatzdaten an das IQTIG nicht möglich.

Ein LE kann unter Umständen mehrere LE-identifizierende Daten (mehrere IKNR/BSNR) besitzen. Dies kann dazu führen, dass bei der Abrechnung mit den Krankenkassen ein anderes LE-identifizierendes Datum verwendet wird als bei der Lieferung der QS-Daten an die LQS/KV. Dadurch würden unterschiedliche Pseudonyme eines Leistungserbringers an die BAS übersendet, was dazu führt, dass zusammengehörende Daten nicht zusammengeführt werden können. Daher wird ein Verfahren benötigt, welches die zusammengehörenden Pseudonyme der BAS bekannt macht.



Achtung

Datenlieferfrist zur Übermittlung der Zusatzdaten an das IQTIG

Die Zusatzdaten sind von den jeweiligen DAS bis spätestens **28. Februar** des Folgejahres an das IQTIG zu übermitteln!

Bei der Registrierung des LE bei der LQS/LKG/KV/KZV werden alle vorhandenen LE-identifizierenden Daten erfasst und als Listen zusammengehöriger IKNR/BSNR in der jeweiligen Landesstelle geführt. Die Listen bestehen aus Listen von zusammengehörigen Pseudonymen, Standortnummern, Gültigkeitszeiträumen der LE-identifizierenden Daten und ähnlichen Metadaten und werden in zwei separaten CSV-Dateien abgebildet.

Die Zusatzdatenverwaltung arbeitet mit zwei CSV-Dateien, Zusatzdaten.csv und Pseudomapping.csv, in denen die LQS/LKG/KV/KZV die LE-identifizierenden Daten erfassen, um sie an die BAS zu übermitteln. Beide Dateien sind obligatorisch, um die Daten an die BAS übermitteln zu können!

Zusatzdaten.csv

In der Datei Zusatzdaten.csv werden die Zusatzdaten zu den einzelnen LE-identifizierenden Daten erfasst. Sie ist die führende Datei. Die Inhalte der Zusatzdaten.csv werden vor der Übertragung an die BAS automatisch mithilfe des aktuellen öffentlichen Schlüssels der jeweiligen LQS/LKG/KV/KZV für jeweils alle aktuell gültigen Pseudonymisierungsverfahren pseudonymisiert, in eine für die Übertragung geeignete Datenstruktur überführt und um die Inhalte der Datei Pseudomapping.csv angereichert. In der Zusatzdaten.csv müssen lediglich die Haupt-LE-identifizierenden Daten vorhanden sein. Weitere LE-identifizierende Daten sind in der Datei Pseudomapping.csv zu erfassen. Sollte ein LE nur ein LE-identifizierendes Datum besitzen, muss dieses nicht in der Pseudomapping.csv enthalten sein.

Beispiel: Zusatzdaten.csv

```
pseudonym;ist HauptPseudonym;Standortnummern;gültig ab;gültig bis;notiz
123456789;1;01,02,03;01.01.2015;;
888888888;0;02;01.01.2000;;
999999999;0;;01.01.2000;31.12.2014;zusammenlegung zweier KHS
234567891;1;01,02;;;
345678912;0;02;;;
456789123;1;01,02;;;
567891234;1;99,01;;;
678912345;0;01;;;
```



Achtung

Die CSV-Datei muss eine Kopfzeile besitzen!

Die Spalten der CSV-Datei besitzen folgende Bedeutung:

Tabelle 6: Spaltenbedeutung der CSV-Datei

Spalte	Beschreibung
pseudonym	Das LE-identifizierende Datum, welches vor der Übermittlung an die BAS automatisch pseudonymisiert wird.
ist HauptPseudonym	<ul style="list-style-type: none"> 1, wenn das Pseudonym dieses LE-identifizierenden Datums (IKNR/BSNR) als Haupt-Pseudonym gekennzeichnet werden soll. sonst 0
Standortnummern	(optional) Die für dieses LE-identifizierende Datum gültigen Standortnummern.

Spalte	Beschreibung
gültig ab	(optional) Sollte dieses LE-identifizierende Datum erst ab einem bestimmten Datum gültig sein, kann dies hier in der Form <dd.MM.yyyy>vermerkt werden.
gültig bis	(optional) Sollte dieses LE-identifizierende Datum nur bis zu einem bestimmten Datum gültig sein, kann dies hier in der Form <dd.MM.yyyy>vermerkt werden.
notiz	(optional) Eine Notiz, welche bei der Fehlersuche helfen könnte, wie z. B. „Verkauf eines Standorts, daher IKNR nur gültig bis“.

Pseudomapping.csv

In der Datei `Pseudomapping.csv` werden pro Zeile die zusammengehörenden LE-identifizierenden Daten (IKNR/BSNR) eines LE unverschlüsselt aufgelistet. Die Haupt-IKNR bzw. Haupt-BSNR steht an erster Stelle, gefolgt von den zusätzlichen IKNR/BSNR (siehe Beispiel: `Pseudomapping.csv`).

Vor der Übertragung der Daten an die BAS werden die IKNR/BSNR durch das Pseudonymisierungsprogramm automatisch mithilfe des aktuellen öffentlichen Schlüssels der jeweiligen LQS/LKG/KV/KZV pseudonymisiert. Hierbei werden jedes Hauptpseudonym und alle zugehörigen Pseudonyme automatisch für jeweils alle aktuell gültigen Pseudonymisierungsverfahren pseudonymisiert und an die BAS versendet.

Beispiel: Pseudomapping.csv

```
123456789;888888888;999999999;;;
234567891;345678912;;;
456789123;;;
567891234;678912345;;;

```

Übermittlung der Pseudonyme an die BAS

Mithilfe des PSP kann die Liste von zusammengehörenden Pseudonymen an die BAS übertragen werden. Für die Zusatzdatenverwaltung bei der BAS stellt diese einen SOAP-Webservice zur Verfügung, über welchen Listen von zusammengehörigen Pseudonymen und Zusatzdaten, wie Standortnummern und Validitätszeiträumen, an die BAS gesendet werden können. Das zur Verfügung gestellte Pseudonymisierungsprogramm kann diesen Webservice bedienen und es den Datenannahmestellen der Leistungserbringer (DAS-LE), LQS, LKG, KV und KZV so erleichtern, die benötigten Daten an die BAS zu liefern.

Der SOAP-Webservice bei der BAS ist gegen Angriffe abgesichert. Jede zur Datenlieferung berechnete Stelle erhält von der BAS einen Benutzernamen und ein Passwort für die Benutzung des SOAP-Webservice. Zudem findet die Übermittlung der Daten verschlüsselt statt. Hierzu wird der gesamte SOAP-ENV: Body mit dem AES-Algorithmus im CBC-Verfahren mit einem PKCS5-Padding verschlüsselt.

Der Schlüsselstring kann mit jeder Version des Verschlüsselungsprogramms ausgetauscht werden.

4 Fehlermeldungen

Im Folgenden werden alle Fehlermeldungen beschrieben, die während der Nutzung des Programms ausgegeben werden können.

Sollte das Programm bei einem Doppelklick nicht starten oder das „Öffnen mit“-Fenster von Windows geöffnet werden, ist möglicherweise die Java-Laufzeitumgebung nicht korrekt installiert. In diesem Fall wenden Sie sich bitte an Ihren Systemadministrator.

Geschweifte Klammern {} in der Spalte „Bedeutung“ von Tabelle 9 werden vom Programm dynamisch sinnvoll ersetzt.

Beispiel:

Bei der Verarbeitung der CSV-Datei {} ist ein Fehler in Zeile {} aufgetreten.

Wird ersetzt durch:

Bei der Verarbeitung der CSV-Datei c:/temp/Muster.csv ist ein Fehler in Zeile 25 aufgetreten.

Tabelle 7: Fehlercodes

Fehler-code	Fehlermeldung	Erläuterung
1000	Unbekannt	
1	Konnte <code>certificate.startdate</code> nicht konvertieren.	<code>certificate.startdate</code> ist im falschen Format angegeben.
2	Konnte keine Schlüsselpaare erzeugen.	
4	Konnte den öffentlichen Schlüssel nicht schreiben.	Auf das angegebene ‚Schlüssel Verzeichnis‘ (<code>registration.dir</code>) konnte nicht schreibend zugegriffen werden.
5	Konnte den PKCS12-Schlüssel nicht schreiben.	Auf das angegebene ‚Schlüssel Verzeichnis‘ (<code>registration.dir</code>) konnte nicht schreibend zugegriffen werden.
6	Konnte den öffentlichen Schlüssel nicht lesen.	
7	Der Schlüssel existiert schon.	-
9	Unbekannter Zertifikatstyp	-

Fehler-code	Fehlermeldung	Erläuterung
10	Die im Schlüssel enthaltenen Metadaten enthalten falsche Angaben zur Datenannahmestelle.	-
12	Die Daten konnten nicht pseudonymisiert werden.	
13	Die Daten konnten nicht depseudonymisiert werden.	
14	Falscher Dateiname für den öffentlichen Schlüssel.	-
15	Registrationsverzeichnis existiert nicht.	-
16	Das angegebene Registrationsverzeichnis ist kein Verzeichnis.	-
17	Fehler beim Erzeugen des privaten Schlüssels.	
18	Fehler beim Laden des privaten Schlüssels. Entweder das Passwort ist falsch oder das Zertifikat ist defekt.	-
19	Der Klartext {} konnte nicht vom Verfahren getrennt werden.	Das gerade depseudonymisierte Pseudonym ist aus einem nicht der Spezifikation entsprechenden Klartext erzeugt worden. Bitte benachrichtigen Sie die BAS über diesen Vorfall.
20	Die Property mit dem Namen {} wird benötigt.	-
21	Die Property mit dem Namen {} muss einen Integer-Wert enthalten.	-
22	CSV-Datei {} konnte nicht gefunden werden.	-
26	Es gibt keine Länderkürzel mit dem Namen oder Kürzel {}.	-
27	Es gibt kein Modul mit dem Kürzel {}.	-
28	Es gibt kein Verfahren mit dem Kürzel {}.	-
29	Es gibt keine Datenannahmestelle mit dem Anfangskürzel {}.	-
30	Keine Schlüssel für die Datenannahme {} {} im Verzeichnis {} gefunden.	-

Fehler-code	Fehlermeldung	Erläuterung
32	In der LQS kann kein Bundesland {} ausgewählt werden.	-
33	In der KV kann kein Bundesland {} ausgewählt werden.	-
34	Die CSV-Datei hat nur {} Spalten in {} wurde {} angegeben.	-
35	Es wurde kein Zertifikat gefunden, welches zum angegebenen Datum {} valid ist.	-
36	Die benutzte Datumsformatierung ist nicht valid {} muss dd.MM.yyyy genügen.	-
37	Das Pseudonym {} konnte nicht von der Zertifikats-ID getrennt werden.	Das angegebene Pseudonym entspricht nicht der Spezifikation. Bitte benachrichtigen Sie die BAS über diesen Vorfall.
38	Die Zertifikats-ID {} ist invalid.	Die angegebene Zertifikats-ID entspricht nicht der Spezifikation. Bitte benachrichtigen Sie die BAS über diesen Vorfall.
39	Das Pseudonym wurde von einer anderen Datenannahmestelle pseudonymisiert. Die Zertifikats-ID der fremden DAS lautet {}.	-
40	Zeile {} in CSV-Datei {} ist nicht korrekt formatiert.	-
41	In der CSV-Datei {} in der Zeile {} fehlt ein Hauptpseudonym.	-
42	Bei der Übertragung der Daten an die BAS ist ein Fehler aufgetreten: {}	
43	Bei der Verarbeitung der CSV-Datei {} ist ein Fehler in Zeile {} aufgetreten.	
1001	Konsolen-Argument 'p', 'd' oder 'z' erforderlich!	-
1002	Konsolen-Argument {} erforderlich!	-

Bei der Verwendung der grafischen Benutzeroberfläche oder des Konsolenprogramms wird im gleichen Verzeichnis eine `pseudonymisierungs.log`-Datei geschrieben. In dieser Datei werden Programmvorgänge und Fehler protokolliert. Die dieses Protokoll beinhaltende Datei ist der Schlüssel zur Fehleranalyse.

Wird die Referenzimplementierung verwendet, sollte das Slf4j-Logging-Framework konfiguriert werden, damit auch hier die entsprechenden Informationen protokolliert werden.

Die meisten möglichen Fehlermeldungen sind selbsterklärend. Wird auf dieses Kapitel verwiesen, macht ein Blick in diese Logdatei Sinn.

Sofern die Fehleranalyse von der Datenannahmestelle nicht allein bewältigt werden kann, sollte diese Datei nach Auftreten des Fehlers gesichert und bei Unterstützungsanfragen an die BAS mitgesendet werden.

Anhang

Tabelle 8: Gültige Länderkürzel

Kürzel	Bedeutung
BU	Bundesweit
BA	Bayern
BB	Brandenburg
BE	Berlin
BW	Baden-Württemberg
HB	Bremen
HE	Hessen
HH	Hamburg
MV	Mecklenburg-Vorpommern
NI	Niedersachsen
NO	Nordrhein
NW	Nordrhein-Westfalen
RP	Rheinland-Pfalz
SH	Schleswig-Holstein
SL	Saarland
SN	Sachsen
ST	Sachsen-Anhalt
TH	Thüringen
WL	Westfalen-Lippe
TESTLAND	Zur Nutzung auf den Testdatenstrecken zwischen Softwareanbietern und Datenannahmestellen

Tabelle 9: Gültige Datenannahmestellen

Kürzel	Datenannahmestelle
LQS	Landesgeschäftsstelle für Qualitätssicherung
BAS	Bundesauswertungsstelle
KV	Kassenärztliche Vereinigung
SV	Datenannahmestelle für selektivvertraglich erbrachte Leistungen