# Anwenderhandbuch


Version 1.13 vom 31.10.2022

*Herausgeber*:

Unabhängige Treuhandstelle der Universitätsmedizin Greifswald

*Autor:*

Christopher Hampf, M.Sc.

Ellernholzstr. 1-2
17475 Greifswald

Tel. 03834 / 86-7851, Fax: 03834 / 86-6843
E-Mail: christopher.hampf@uni-greifswald.de

Unabhängige Treuhandstelle
UNIVERSITÄTSMEDIZIN GREIFSWALD

# Versionierung

| *Version* | Datum | Bearbeitungsart / Betroffene Abschnitte | Bearbeiter |
|---|---|---|---|
| *1.0* | 25.03.2019 | Update und Erweiterung der Doku von mosaic-greifswald.de | Christopher Hampf |
| *1.0.1* | 19.06.2019 | Hinzufügen des Hinweises auf etwaige Probleme hinsichtlich der Verwendung von Volumes und VPN-Clients | Christopher Hampf |
| *1.0.2* | 15.07.2019 | Anpassung von Begrifflichkeiten | Christopher Hampf |
| *1.0.3* | 09.08.2019 | Domänenschnittstelle ergänzt | Martin Bialke |
| *1.9.0* | 21.11.2019 | Abgleich der Versionierung mit gPAS<br>Diverse Ergänzungen und Korrekturen | Christopher Hampf |
| *1.9.1* | 04.03.2020 | Web-Auth für Version 1.9.1 ergänzt | Martin Bialke |
| *1.10.2* | 11.03.2021 | Aktualisierung der Bilder zu Benutzeroberflächen<br>Ergänzungen zu neuen Funktionalitäten<br>Autorisierung/Authentifizierung überarbeitet<br>FHIR Gateway ergänzt<br>Optimierungen ergänzt<br>Logging ergänzt | Martin Bialke<br>Christopher Hampf |
| *1.10.2* | 12.07.2021 | Link zu https://www.ths-greifswald.de/ttp-tools/keycloak ergänzt und Links FHIR GW aktualisiert | Martin Bialke |
| *1.11.0* | 15.11.2021 | Installationsanleitung angepasst<br>Aktualisierung der Bilder zu Benutzeroberflächen<br>Upgrade ergänzt | Christopher Hampf |
| *1.12.0* | 31.03.2022 | Aktualisierung auf gPAS Version 1.12.0 | Christopher Hampf und Martin Bialke |
| *1.12.1* | 30.06.2022 | Hinweis ergänzt zur Nutzung von KeyCloak in Kombination mit HTTP2 | Christopher Hampf |

| | | | |
|---|---|---|---|
| *1.13* | 31.10.2022 | Aktualisierungen auf gPAS 1.13 | Christopher Hampf |

# Inhalt

# Abbildungsverzeichnis

# Tabellenverzeichnis

# 1 Hintergrund

Die Durchführung klinisch-epidemiologischer Studien, aber auch der Aufbau von Registern und Kohorten, erfordern eine datenschutzkonforme Datenverarbeitung. Gemäß Art. 32 Abs. 1a DSGVO unterstützt die Verwendung von Pseudonymen dabei, ein angemessenes Schutzniveau der Datenverarbeitung zu gewährleisten. Am Institut für Community Medicine der Universitätsmedizin Greifswald wurde hierfür der Generic Pseudonym Administration Service (kurz: gPAS®) entwickelt. Das Web-service-basierte Werkzeug gPAS® dient der Generierung und Verwaltung von Pseudonymen. Das Domänenkonzept sowie die freie Definition von Alphabeten als auch Generatoralgorithmen erlauben unterschiedliche Pseudonyme je Datenquelle, Anwendungskontext (z.B. Erhebung oder Herausgabe) und Standort zu generieren.

Der gPAS® ist als Open Source Software lizensiert (AGPLv3) und kostenfrei für kommerzielle und nicht-kommerzielle Zwecke einsetzbar.

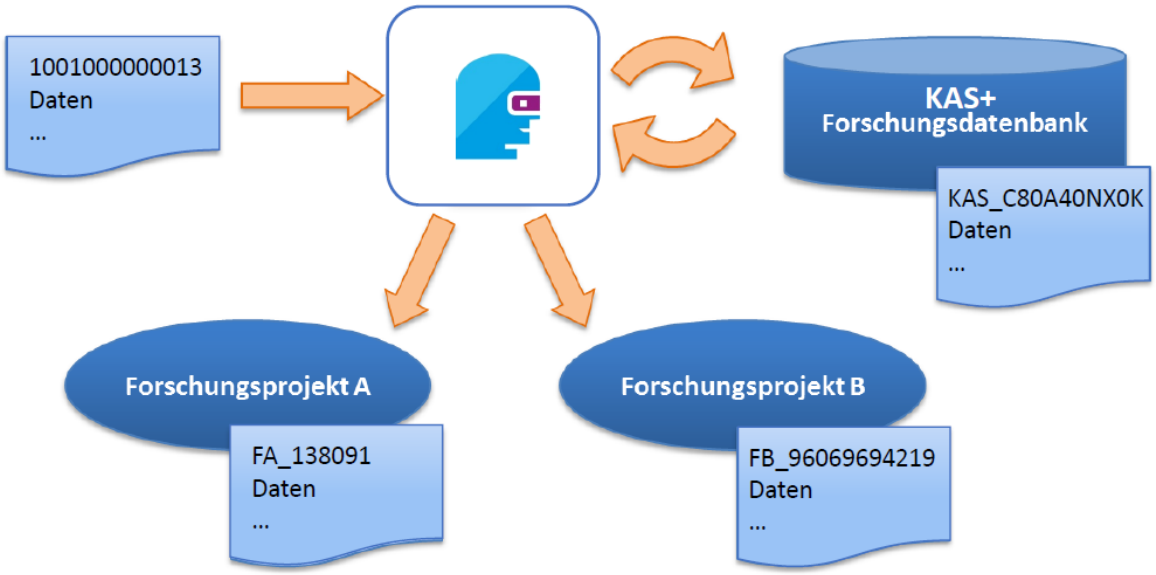



**Abbildung 1-1:** Zuweisung und Verwaltung von projektspezifischen Pseudonymen zu einem Personenidentifikator mittels gPAS®.

# 2 Begriffsbestimmungen

### Domäne (Domain)

Eine Domäne ist eine organisatorische Einheit (Mandant), z.B. eine Studie, ein Projekt oder ein Institut.

### Originalwert

Der Originalwert bezeichnet den Wert, für den ein Pseudonym höherer Stufe generiert und diesem Originalwert im Anschluss zugewiesen wird.

### Pseudonym

Nicht-sprechender Identifikator, welcher einer Person zugewiesen ist. Ein Pseudonym erster Stufe verweist dabei direkt auf die Personendaten, ein Pseudonym höherer Stufe jeweils nur auf das Pseudonym der jeweils niedrigeren Stufe.

#### Pseudonymisieren

Erzeugen eines nicht-sprechenden Identifikators, basierend auf einem gegebenen Originalwert.

#### Depseudonymisieren

Ermitteln des Originalwerts, eines gegebenen Pseudonyms.

#### Anonymisieren

Veränderung von Patientendaten zur jeweiligen natürlichen Person, dass eine Zuordnung nicht oder nur mit Verhältnismäßig viel Aufwand möglich ist. Im gPAS® werden keine Personendaten gespeichert, weshalb eine Anonymisierung das unwiederbringliche Löschen von Zuordnungen zwischen originalwert und Pseudonymen vorsieht. Ein Rückschluss auf die Person ist dann nicht mehr möglich.

# 3 Funktionalitäten

## 3.1 Was leistet der Dienst

- Generierung von Pseudonymen (PSN)
- Zuordnung von Pseudonymen zu beliebigen Originalwerten
- Technische Anonymisierung durch Löschung von Zuordnungen zwischen Pseudonym und Originalwert
- Konfiguration von Pseudonym-Parametern: Prüfzifferalgorithmus, Länge, Alphabet
- Verwaltung von Pseudonym-Domänen und Zweitpseudonymen
- Validierung von Pseudonymen
- Depseudonymisierung
- Darstellung von Pseudonym-Hierarchien
- Import und Export vorhandener Pseudonyme
- Löschen von temporären Pseudonymen (dabei werden sowohl Pseudonym als auch zugeordneter Originalwert gelöscht)
- Listenverarbeitung
- Hohe Performance durch Caching
- Unterstützung für KeyCloak-Authentifizierung und Autorisierung

## 3.2 Was leistet der Dienst nicht

- Maskierung: der erforderliche Schritt zur Trennung von personenidentifizierenden und medizinischen Daten ist nicht Teil des gPAS®-Systems und muss vom nutzenden Projekt geleistet werden
- Extraktion und Schwärzung identifizierender Merkmale in Dokumenten oder Datensätzen
- Record Linkage / Identitäts-Matching[1]

# 4 Installation

## 4.1 Systemanforderungen

### Technisch / Infrastruktur

- Installierte aktuelle Version von Docker[2] und Docker-Compose[3]
- Administrative Rechte
- Keine Nutzungsbeschränkungen auf die bereitgestellten Service- und Client-URLs
- Windows oder Ubuntu Server (oder vergleichbar) mit min. 8 GB Arbeitsspeicher, 5 GB Festplattenspeicher, Prozessor (benötigter Arbeitsspeicher und Prozessor-Leistung sind abhängig von erwarteter Datenmenge und -durchsatz)

### Software: Anwendungs- und Datenbankserver (ohne Verwendung von Docker)[4]

- JDK 17 oder höher
- WildFly 26 oder höher
- EclipseLink 2.7.11MySQL-Connector 8 oder höher
- MySQL-Server 8 oder höher

### Personell

- Verantwortlicher mit grundlegenden IT-Kenntnissen zur Administration des Servers und zur Einrichtung des gPAS®-Dienstes (zuzüglich der Wartung und regelmäßiger Sicherungen der gPAS®-Datenbank)
- Verantwortlicher zur Administration und Pflege der gPAS®-Inhalte

[1] Nutzen Sie hierzu bitte ggf. den E-PIX: http://ths-greifswald.de/e-pix
[2] Weitere Informationen unter https://docs.docker.com/install/
[3] Weitere Informationen unter https://docs.docker.com/compose/install/
[4] Beim Betrieb unter Windows ist zu beachten, dass bei der Verwendung von Volumes und parallel betriebenen VPN-Clients Probleme auftreten können.

## 4.2 Download und Starten des Dienstes

⚠ **Hinweis:** Die hier beschriebene Installation erfolgt standardmäßig mit *Docker-Compose*. Wenn der gPAS® davon abweichend ohne Docker betrieben werden soll, kann eine Performancesteigerung erreicht werden, indem die Hinweise in **Kapitel 12** berücksichtigt werden. In der ausgelieferten Docker-Variante sind diese bereits berücksichtigt und es sind keine weiteren Anpassungen erforderlich.

Um den gPAS® als Docker-Container zu starten, werden die Programme *Docker* und *Docker-Compose* benötigt. Beide Programme müssen hierfür installiert sein. Da zwischen beiden Programmen Inkompatibilitäten auftreten können, wird empfohlen die jeweils aktuellsten Versionen zu installieren.

Der gPAS® benötigt zur Ausführung mehrere Container (vgl. **Abbildung 4-1**). Damit diese nicht einzeln gestartet werden müssen und entsprechend zusammengeschaltet werden müssen, wird der Dienst mit *Docker-Compose* gestartet. Die entsprechenden Ressourcen können von der THS-Webseite heruntergeladen werden[5].

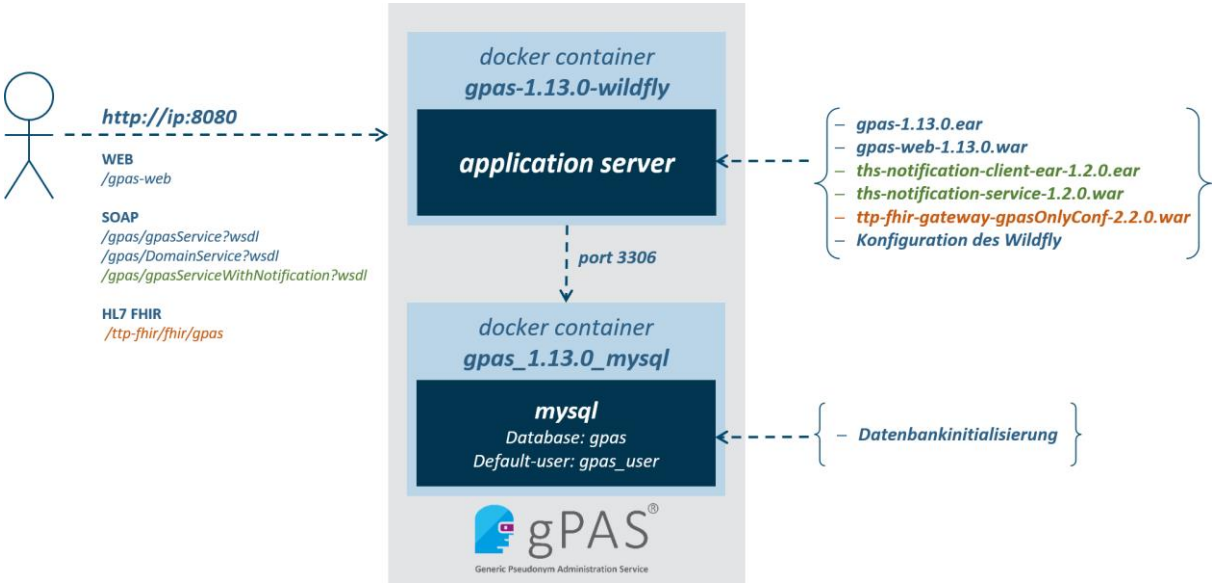



**Abbildung 4-1:** Architektur des gPAS® mit Docker.

Das Docker-System besteht aus zwei getrennten Containern. Zum einen aus einer Datenbankinstanz (MySQL) und zum anderen aus dem Anwendungsserver (WildFly inkl. Datenbank-Konnektoren). Der Anwendungsserver kommuniziert mit dem MySQL-Server über den Port 3306. Der Zugriff auf das System von „außen“ erfolgt über den Web-Browser. Die Inhalte werden über den Port 8080 (gPAS) für den Anwender bereitgestellt.

Um die folgenden Schritte problemlos durchführen zu können, wird ein Account mit administrativen Rechten benötigt. Exemplarisch werden die folgenden Befehle mit `sudo` ausgeführt.

[5] https://www.ths-greifswald.de/forscher/gpas/ bzw. https://www.ths-greifswald.de/forscher/gpas/#download

### Download der benötigten Dateien

Laden Sie die aktuellste Version von https://www.ths-greifswald.de/forscher/gpas/#download herunter und entpacken Sie die ZIP-Datei. Diese enthält alle relevanten Docker-Compose-Dateien. Im Folgenden wird davon ausgegangen, dass der Ordner in das Verzeichnis `/opt/` entpackt wurde. Der Pfad kann bei Bedarf angepasst werden.

### Vergabe von Schreibrechten

```
sudo chmod -R 755 /opt/compose-wildfly/

chown -R 1000:1000 /opt/compose-wildfly/logs/ /opt/compose-wildfly/deployments/
```

Aus Gründen von Leistung und Ausfallsicherheit sollten die Container des gPAS auf einem dedizierten Server eingerichtet werden. Zur Administration werden der User `gpas` (uid 1000) aus der Gruppe `users` (gid 1000) genutzt.

### Wechseln in das gPAS-Verzeichnis für die Standard-Version

```
cd /opt/compose-wildfly/
```

### Starten des gPAS mithilfe von Docker Compose

```
sudo docker-compose up
```

Damit werden die benötigten Komponenten heruntergeladen[6] und die Konfiguration von MySQL und WildFly gestartet. Danach wird die aktuelle Version des gPAS bereitgestellt. Der Installationsvorgang kann in Abhängigkeit der vorhandenen Internetverbindung etwa 5 Minuten dauern. Der erfolgreiche Start des Dienstes wird mit der folgenden Ausgabe abgeschlossen.

```
Wildfly 26.1.2.Final […] started in …
```

⚠ **Hinweis:** Weitere Details zur Nutzung von *Docker-Compose* und gPAS® sind der beigelegten Beschreibung docker-compose/README.md zu entnehmen.

⚠ Hinweis: Beachten Sie, dass beim Wechsel vom gPAS Version 1.x auf 1.13.0 auch die Docker-Compose Komponenten komplett aktualisiert werden (Umstieg auf Java 17 und WildFly 26). Weitere Hinweise zur Aktualisierung der Docker-Komponenten können der beigelegten Docker-Update.md entnommen werden.

[6] Sollte Ihre Maschine keinen Zugang zum Internet haben, können die benötigten Images (MySQL und Wildfly) von einer anderen Maschine heruntergeladen werden und dann auf Ihr Zielsystem kopiert werden (siehe https://docs.docker.com/engine/reference/commandline/image_save/ und https://docs.docker.com/engine/reference/commandline/load/).

# 5 Die grafische Benutzeroberfläche des gPAS®

Um dem Datentreuhänder die Administration der Pseudonyme zu ermöglichen, verfügt der gPAS® über eine grafische Benutzeroberfläche die speziell für den Einsatz im Web-Browser entwickelt wurde.

Der Aufbau der Oberfläche orientiert sich an typischen Arbeitsabläufen eines Datentreuhänders.

## 5.1 Anwendungsfall 1: Anlegen einer Domäne

Um ein neues Pseudonym zu erstellen, muss vorab eine entsprechende **Domäne** vorhanden sein.

Dabei gilt:

- ein Pseudonym ist innerhalb einer Pseudonym-Domäne eineindeutig
- eine Domäne kann für ein Projekt oder ein eingebettetes Studienvorhaben stehen, aber auch zur Beherbergung von Zweitpseudonymen angelegt werden
- für jede Domäne lassen sich eigene Pseudonym-Parameter festlegen

Unter dem Menüpunkt *Domänen* werden alle bereits angelegten Domänen als Baum-Struktur dargestellt (vgl. **Abbildung 5-1**). Mit einem Rechtsklick auf einen Eintrag werden weitere Optionen zur jeweiligen Domäne angeboten. So kann beispielsweise direkt die Seite *Suchen / Anlegen* aufgerufen werden, wobei die jeweilige Domäne vorausgewählt ist. Um eine neue Domäne zu erstellen, kann entweder die Schaltfläche *Erstellen* angewählt werden oder über das Kontextmenü je nach angewählter Domäne eine *Geschwister-* oder eine *Kind-Domäne* angelegt werden. Daraufhin öffnet sich ein Fenster mit den auszufüllenden Feldern. Bei der Wahl einer *Geschwister-Domäne* wird die *Eltern-Domäne* der ausgewählten Domäne übernommen. Bei einer *Kind-Domäne* wird die angewählte Domäne als *Eltern-Domäne* gesetzt.

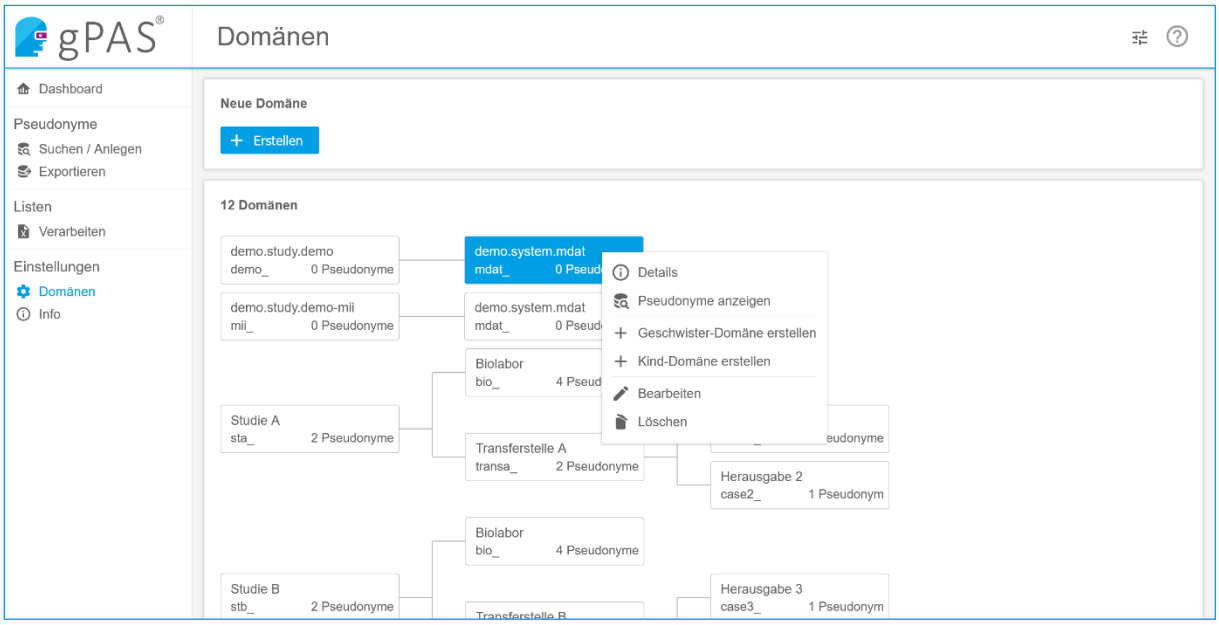



**Abbildung 5-1:** Oberfläche zum Anzeigen aller Domänen. Der Baum zeigt die hierarchische Struktur der Domänen. Mit einem Rechtsklick auf eine Domäne öffnet sich das Kontextmenü, welches weitere Optionen enthält.

Beim Anlegen einer neuen Domäne ist der eindeutige Name festzulegen. Optional kann eine oder mehrere *Eltern-Domänen* (um Pseudonym-Hierarchien zu realisieren) angegeben werden. Eine Domäne mit mehreren *Eltern-Domänen* wird entsprechend oft in den jeweiligen Domänen-Bäumen dargestellt. Zusätzlich kann eine Beschreibung ergänzt werden.

Das zu nutzende Alphabet (vgl. **Tabelle 5-1**) und der jeweilige Prüfziffern-Generator (vgl. **Tabelle 5-2**) können in sinnvollen Kombinationen ausgewählt werden. Nur erlaubte Kombinationen werden dabei angeboten. Die Oberfläche zum Anlegen einer neuen Domäne ist in **Abbildung 5-2** dargestellt.

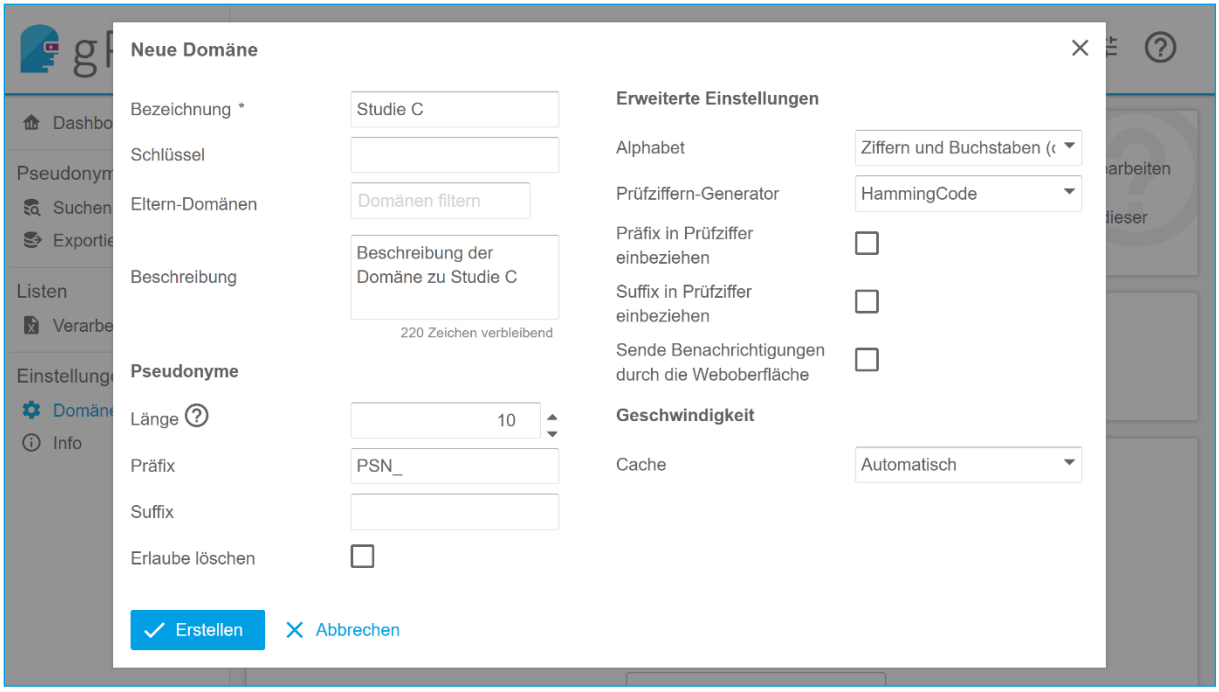



**Abbildung 5-2:** Oberfläche zum Anlegen einer neuen Domäne.

Mögliche Alphabete sind in **Tabelle 5-1** aufgelistet. Alternativ kann ein eigenes Alphabet (aus beliebigen Zeichen bis auf Kommas und Leerzeichen) definiert werden. Hierzu wählt man im Alphabet Auswahlmenü den Eintrag *Benutzerdefiniert* und trägt die gewünschten Zeichen in das erscheinende Textfeld ein (vgl. **Abbildung 5-2**).

**Tabelle 5-1:** Bereitgestellte Alphabete im gPAS®.

| Alphabet | Beschreibung |
|---|---|
| Hex | 16 Zeichen: 0-9, A-F |
| Numbers | 10 Zeichen: 0-9 |
| NumbersWithoutZero | 9 Zeichen: 1-9 |
| NumbersX | 11 Zeichen: 0-9, X |
| Symbols32 | 32 Zeichen: 0-9, A-Z (ohne B, I, O, S – wegen der Ähnlichkeit zu 8, 1, 0, 5) |
| Symbols31 | Wie Symbol32, nur ohne V – wegen der Ähnlichkeit zu U |

Dabei ist zu beachten, dass einzelne Prüfzeichenalgorithmen nur unter bestimmten Bedingungen angewendet werden können. Die entsprechenden Bedingungen der jeweiligen Prüfziffernalgorithmen sind in **Tabelle 5-2** dargestellt. Sollen keine Prüfziffern enthalten sein, wird die Option *Keine Prüfziffern* angewählt.

**Tabelle 5-2:** Mögliche Prüfziffern-Generatoren und dessen Bedingungen.

| Prüfziffernalgorithmus | Bedingung |
|---|---|
| HammingCode | Alphabet-Länge ist eine Primzahlpotenz (Achtung: Derzeitige Implementierung erlaubt nur den Wert 32) |
| Verhoeff | Alphabet-Länge ist gleich 10 |
| VerhoeffGumm | Alphabet-Länge ist gleich 10 |
| Damm | Alphabet-Länge ist gleich 10 |
| Reed-Solomon-Lagrange | Alphabet-Länge ist eine Primzahl, wobei die maximale Anzahl der Prüfzeichen gleich der Alphabet-Länge ist |

Ergänzend können bei jeder Domäne noch weitere Eigenschaften definiert werden. Diese sind mit entsprechenden Beispielen in **Tabelle 5-3** dargestellt.

**Tabelle 5-3:** Mögliche Domain Properties.

| Feld (SOAP-Parameter) | Bedeutung | Beispiel |
|---|---|---|
| Anzahl detektierbarer Fehler (*MAX_DETECTED_ERRORS*) | Max. Anzahl fehlerhafter Zeichen, die auch sicher als fehlerhaft erkannt werden (nur bei Reed-Solomon) | 2 |
| Länge (*PSN_LENGTH*) | Länge des erzeugten Pseudonyms ohne Prüfzeichen (max. 49) | 10 |
| Präfix (*PSN_PREFIX*) | Zeichenfolge, die als Präfix genutzt werden soll (max. 20 Zeichen) | abc_ |
| Suffix (*PSN_SUFFIX*) | Zeichenfolge, die als Suffix genutzt werden soll (max. 20 Zeichen) | _xyz |
| *Präfix in Prüfziffernberechnung einbeziehen* (*INCLUDE_PREFIX_IN_CHECK_DIGIT_CALCULATION*) | Einschließen des Präfix in die Berechnung der Prüfziffern | ja (`true`)/ nein (`false`) |
| Erlaube löschen (*PSNS_DELETABLE*) | Pseudonyme innerhalb der Domäne sind löschbar (z.B. zur Erstellung von | ja (`true`)/ nein (`false`) |

| | | |
|---|---|---|
| | temporären Pseudonymen, der Standard-Wert ist nein(`false`).) | |
| Pseudonym Cache (*FORCE_CACHE*) | ON = Immer an<br>OFF = Immer aus<br>DEFAULT = An, wenn Alphabet + PSN Länge klein genug, dass 120 MB Cache unterschritten werden.[7] | ON, OFF, Automatisch (DEFAULT) |
| *Abstand für Delimiter bzw. Trennzeichen* (*useLastCharAs DelimiterAfterXChars*) | Zahlenwert im Intervall [0, 50[.<br>Wenn größer 0, dann gibt dies den Abstand an, in der das Trennzeichen (letztes Zeichen im benutzerdefinierten Alphabet) im Pseudonym gesetzt wird. Bei 0 wird kein Trennzeichen gesetzt. | Wenn z.B. der Wert 3 ist:<br>123-456-789<br>(„-“ ist das Trennzeichen) |

Das **Löschen einer Domäne** ist nur möglich, sofern keine Pseudonyme für diese Domäne hinterlegt sind.

Hingegen ist das **Löschen eines Pseudonyms** möglich, wenn die Domänenkonfiguration dies gestattet (*Erlaube löschen* unter dem Punkt *Pseudonyme* (`PSNS_DELETABLE=TRUE`)).

Die Geschwindigkeit bei großen Beständen kann erheblich verbessert werden, wenn das integrierte **Caching** verwendet wird. Standardmäßig aktiviert der gPAS® dies bei sinnvollen Beständen selbst (Option *automatisch*). Es kann jedoch expliziert für eine Domäne aktiviert oder deaktiviert werden.

Darüber hinaus kann der gPAS® Notifications versenden, um andere Systeme über Änderungen zu Informieren. Weitere Informationen können dem Kapitel  entnommen werden.

## 5.2 Anwendungsfall 2: Generieren von Pseudonymen

### Generieren eines einzelnen Pseudonyms

Unter dem Menüpunkt *Suchen / Anlegen* können einzelne Pseudonyme angelegt werden. Hierzu muss zunächst die Schaltfläche *Neues Pseudonym generieren* angewählt werden. Im sich öffnenden Fenster muss zum einen die Domäne und zum anderen der Originalwert angegeben werden. Ist der Originalwert beispielsweise ein Pseudonym erster Stufe, wird ein Pseudonym zweiter Stufe generiert. Das generierte Pseudonym folgt dann den entsprechenden Vorgaben der Domäne (siehe vorheriger Abschnitt). In **Abbildung 5-3** wird die Oberfläche zum Anlegen eines Pseudonyms dargestellt.

[7] Memory consumption is one bit per possible pseudonym: mem_for_cache = alphabet_length ^ pseudonym_length / 8 / 1024 / 1024 MB, e.g. alphabet = numbers, length = 8 -> mem_for_cache = 10 ^ 8 / (8 * 1024 * 1024) = 11.92 MB

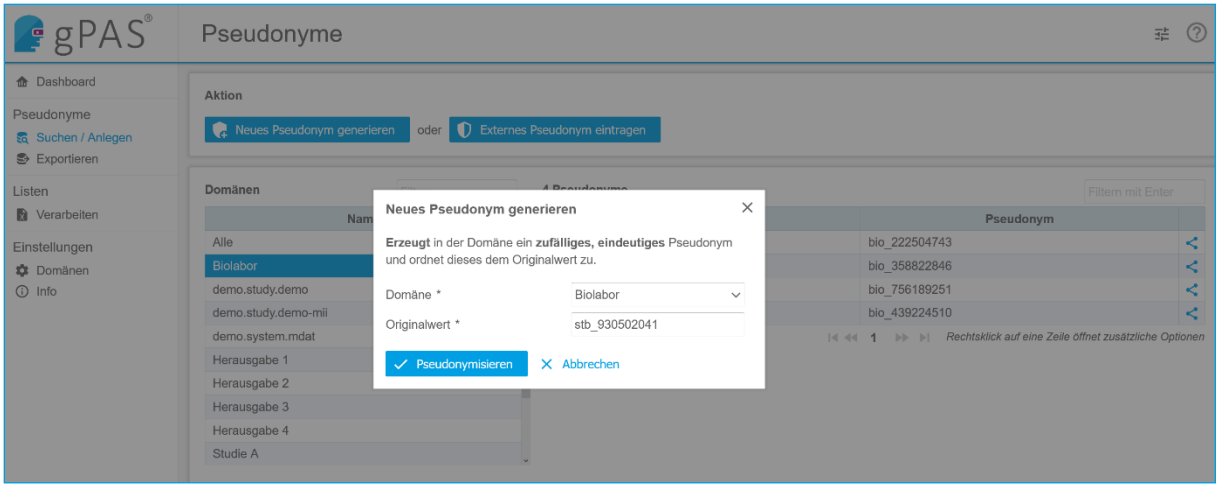



**Abbildung 5-3:** Oberfläche zum Anlegen eines neuen Pseudonyms.

## Generieren eines Zweit-Pseudonyms

Oftmals ist das Generieren von Zweit-Pseudonymen (oder beliebig vielen Pseudonymen) erforderlich, z.B. bei unterschiedlichen Pseudonymen je Studienzentrum und Datentyp. In diesem Fall empfiehlt es sich jeweils eigene Domänen anzulegen. Als Originalwerte werden dann die bei der Erst-Pseudonymisierung generierten Pseudonyme bzw. die Pseudonyme der niedrigeren Stufe verwendet.

Bei einem Pseudonym höherer Stufe wird als Originalwert ein zuvor generiertes Pseudonym verwendet. Um ein Zweit-Pseudonym zu generieren, kann wie eben beschrieben vorgegangen oder das Kontextmenü per Rechtsklick genutzt werden. Über die Schaltfläche *Pseudonymisiere Originalwert* wird ein Pseudonym derselben Stufe generiert. Über die Schaltfläche *Pseudonymisiere Pseudonym* wird das Pseudonym als Originalwert verwendet und es wird ein Pseudonym höherer Stufe erzeugt. In **Abbildung 5-4** wird das entsprechende Kontextmenü dargestellt.

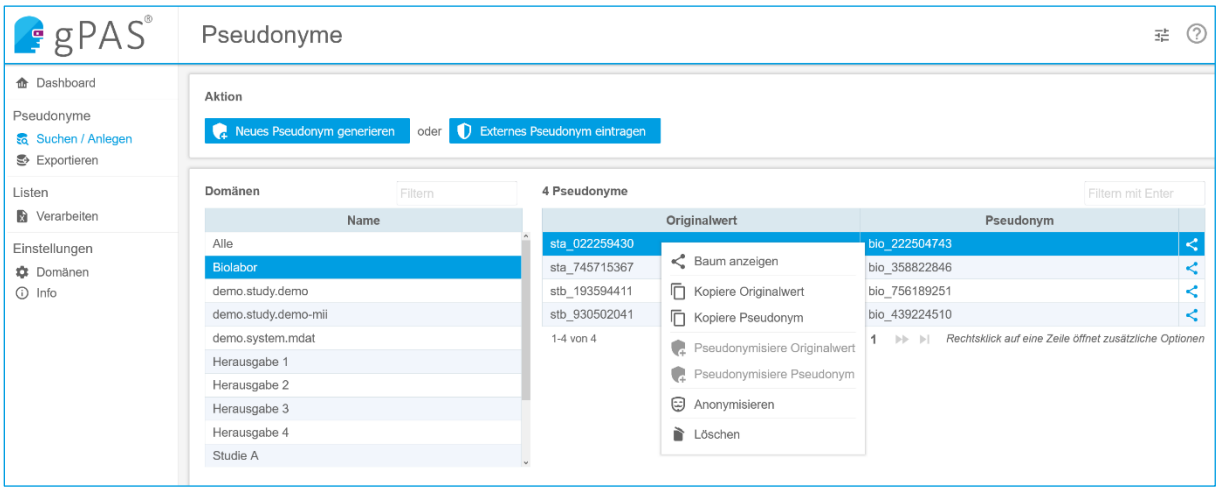



**Abbildung 5-4:** Kontextmenü zum Erzeugen von Pseudonymen derselben Stufe (*Pseudonymisiere Originalwert*) und einer höheren Stufe in einer Kind-Domäne (*Pseudonymisiere Pseudonym*).

In **Abbildung 5-5** wird eine beispielhafte Struktur für mehrere Pseudonyme mit mehreren Stufen dargestellt. Der Studienteilnehmer hätte in diesem Fall zwei Pseudonyme zweiter Stufe, jeweils eins für Studie A und eins für Studie B. Bei beiden stellt das Pseudonym erster Stufe (1001000000011) den Originalwert dar. Dies könnte beispielsweise ebenfalls ein Master Patient Index aus dem E-PIX® sein.

Basierend auf dem Pseudonym für Studie B, wurden zwei weitere Pseudonyme dritter Stufe generiert. Der Originalwert der Pseudonyme dritter Stufe ist dabei das Pseudonym zweiter Stufe.

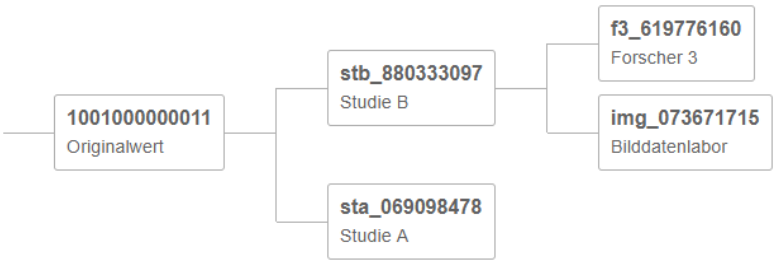



**Abbildung 5-5:** Exemplarische Struktur bei mehreren Pseudonymen und Stufen für einen Studienteilnehmer.

## 5.3 Anwendungsfall 3: Originalwerte und Pseudonyme suchen

Originalwerte und Pseudonyme werden auf die gleiche Weise gesucht. Hierzu wird unter dem Menüpunkt *Suchen / Anlegen* eine Domäne angewählt. Wenn ein Originalwert gesucht wird, muss hierzu die Domäne gewählt werden, in der sich das höherstufige Pseudonym befindet. Wenn ein Pseudonym gesucht wird, dann wird die Domäne gewählt, in der sich das jeweilige Pseudonym befindet. In das obere rechte Suchfeld (über der Pseudonym-Auflistung) wird die gesuchte Zeichenkette eingetragen. Die dargestellte Tabelle wird nach dem Drücken der Enter-Taste gefiltert. Es werden nur die Einträge angezeigt, bei der eine exakte Übereinstimmung vorhanden ist. Hierbei ist es irrelevant an welcher Position sich die eingegebene Zeichenkette befindet. In **Abbildung 5-6** ist die Oberfläche zum Suchen von Originalwerten und Pseudonymen dargestellt.

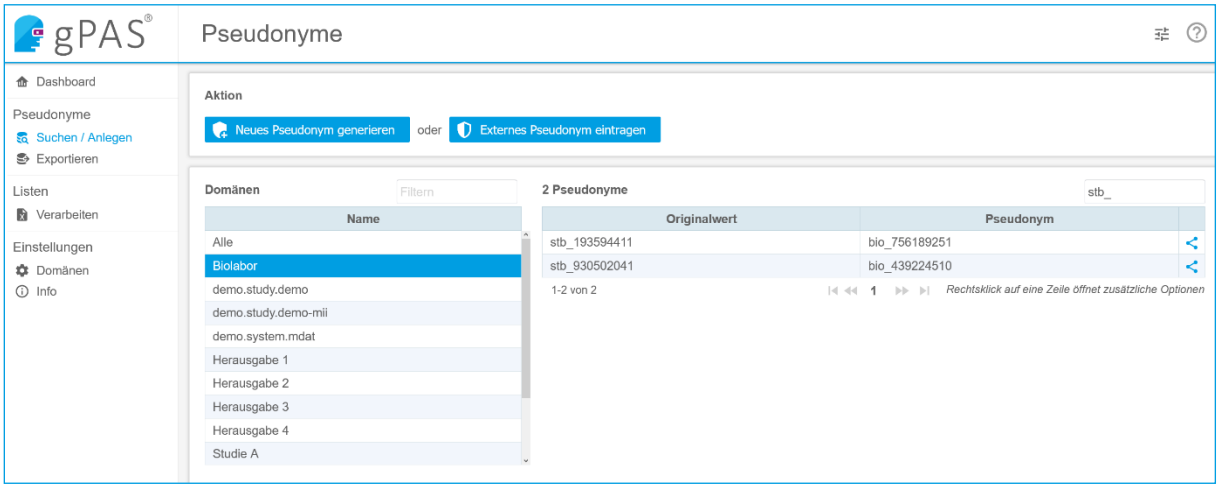



**Abbildung 5-6:** Oberfläche zum Suchen von Originalwerten oder Pseudonymen.

Eine Suche über alle Domänen hinweg, kann durchgeführt werden, indem in der Domänenauflistung der Eintrag *Alle* gewählt wird. Die Suche erfolgt dann wie eben beschrieben über das Suchfeld. Die gefilterte Tabelle beinhaltet zudem noch die Angabe, aus welcher Domäne der jeweilige Eintrag stammt.

## 5.4 Anwendungsfall 4: Suchen von Originalwerten (Depseudonymisierung)

Eine Depseudonymisierung entspricht der Suche eines Originalwerts anhand eines gegebenen Pseudonyms. Es wird demnach so vorgegangen, wie im **Abschnitt 5.3** beschrieben. Hierzu wird das vorhandene Pseudonym in der jeweiligen Domäne gesucht. Die Suche liefert den dazugehörigen Originalwert bzw. im Fall einer Pseudonym-Hierarchie das Pseudonym geringerer Stufe.

## 5.5 Anwendungsfall 5: Technische Anonymisierung (Virtuelle Anonymisierung)

Bei der technischen Anonymisierung (bekannt auch als virtuelle Anonymisierung und nachfolgend kurz als Anonymisierung bezeichnet) wird die Zuordnung zwischen einem Originalwert und einem Pseudonym **unwiederbringlich** aufgehoben. Das Pseudonym bleibt dabei erhalten, der Originalwert hingegen wird durch einen neu generierten Platzhalter ersetzt. Zur Anonymisierung wird hierzu unter dem Menüpunkt *Suchen / Anlegen* die entsprechende Domäne gewählt. Danach wird in der Liste das zu anonymisierende Pseudonym gewählt (zuvor kann mittels Suche die Liste gefiltert werden, vgl. **Abschnitt 5.3**). Mit einem Rechtsklick auf den entsprechenden Eintrag wird das Kontextmenü geöffnet (siehe **Abbildung 5-7**).

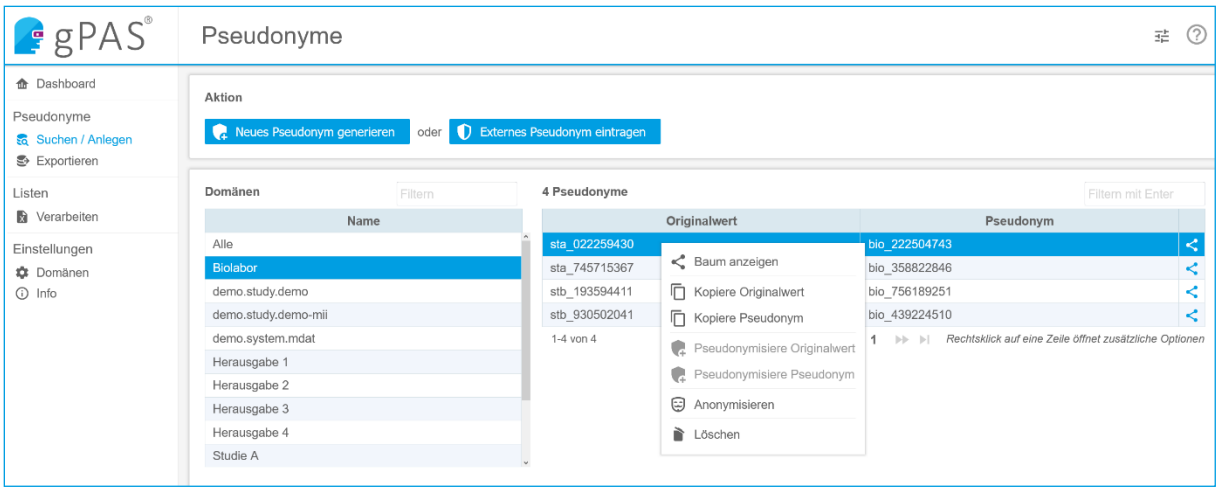



**Abbildung 5-7:** Mit einem Rechtsklick auf den Eintrag, kann das Kontextmenü aufgerufen werden.

Mit dem Auswählen des Eintrags *Anonymisieren*, wird die Zuordnung aufgehoben. Alternativ dazu, kann bei Betrachtung der Baumstruktur (vgl. **Abbildung 5-11**) die Verbindung mittels des kleinen Scheren-Symbols aufgelöst werden (siehe **Abbildung 5-8**). In jedem Fall, muss die Durchführung dieses unumkehrbaren Vorgangs bestätigt werden (siehe **Abbildung 5-9**).

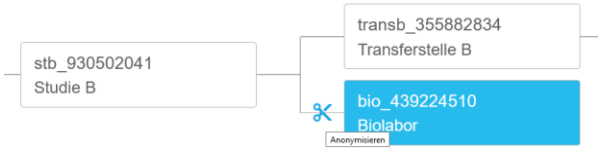



**Abbildung 5-8:** Anonymisierung in der Baumstruktur durch Auftrennen der Verbindung (Schere).

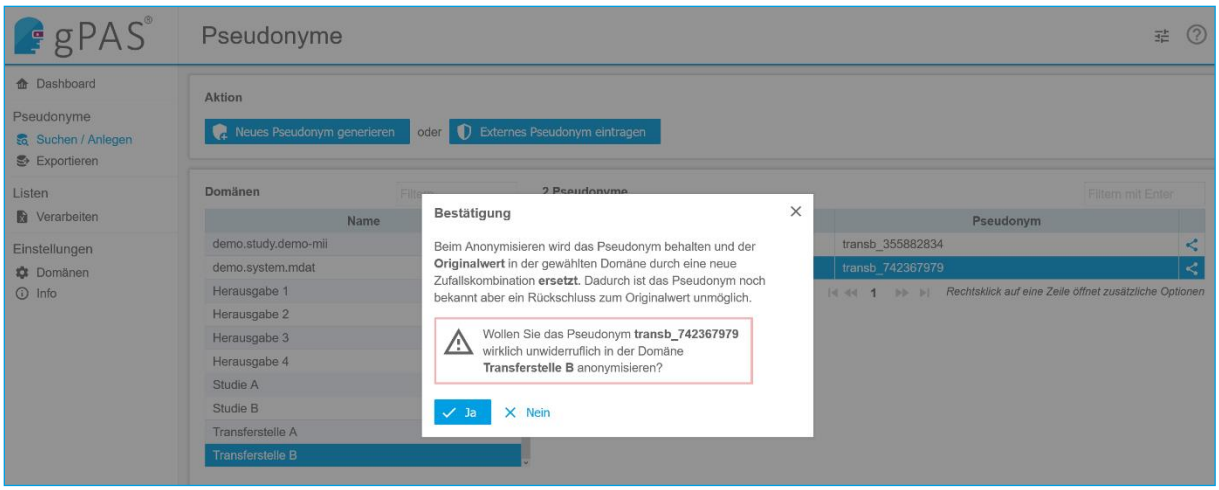



**Abbildung 5-9:** Der unumkehrbare Vorgang der Anonymisierung muss nochmals bestätigt werden.

Nach Bestätigung des Vorgangs, wird die Anonymisierung durchgeführt und der Eintrag entsprechend aktualisiert. Die Anonymisierung ist dadurch ersichtlich, dass der Originalwert mit dem Wert *###_anonym_###_..._###_anonym_###* ersetzt wird (siehe **Abbildung 5-10**).

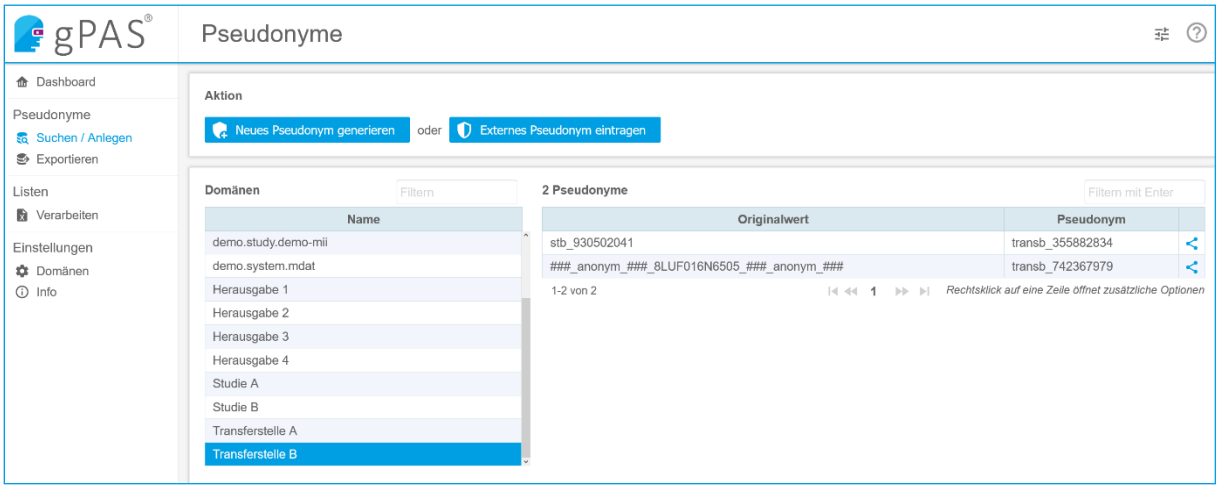



**Abbildung 5-10:** Anonymisierter Eintrag.

## 5.6 Anwendungsfall 6: Löschen von Pseudonymen

Beim Löschen eines Pseudonym-Paares, wird der Originalwert und das entsprechende Pseudonym unwiederbringlich aus der Domäne entfernt.

War das Pseudonym Teil einer Hierarchie bleiben die anderen Elemente der Hierarchie erhalten und müssen bei Bedarf ebenfalls gelöscht werden. In einer niedrigeren Hierarchiestufe bleibt der gelöschte Originalwert also als Pseudonym erhalten. In einer höheren Stufe bleibt das gelöschte Pseudonym als Originalwert erhalten.

Eine Depseudonymisierung ist mittels des gelöschten Pseudonyms nicht mehr möglich. Um ein Pseudonym löschen zu können, muss eine Domäne das zulassen und entsprechend konfiguriert sein (vgl. **Tabelle 5-3** in **Abschnitt 5.1**). Zum Löschen wird unter dem Menüpunkt *Suchen / Anlegen* die entsprechende Domäne und das zu löschende Pseudonym angewählt. Mit einem Rechtsklick auf den

Eintrag wird das Kontextmenü geöffnet (siehe **Abbildung 5-7**). Mit dem Auswählen des Eintrags *Löschen*, wird der Eintrag gelöscht. Bevor dieser unumkehrbare Vorgang ausgeführt wird, muss dies bestätigt werden.

## 5.7 Anwendungsfall 7: Integration von Alt-Pseudonymen

Um ein Pseudonym anzulegen, welches als Originalwert eines bereits existenten Pseudonyms dient, muss unter dem Menüpunkt *Suchen / Anlegen* die Schaltfläche *Von Extern eintragen* gewählt werden. In das sich öffnende Fenster kann die Domäne, der Originalwert und das Pseudonym eingetragen werden. Dabei bezieht sich die Domäne auf das anzulegende Pseudonym. Es ist nicht möglich, ein bereits existierendes Pseudonym erneut anzulegen.

## 5.8 Anwendungsfall 8: Anzeige von Pseudonym-Hierarchien

Wurden bei der Konfiguration der Pseudonym-Domäne eine Eltern-Domäne zur Angabe der Beziehung zweier Domänen (Eltern-Kind-Beziehung) angegeben, so kann der dadurch entstehende Pseudonym-Baum übersichtlich dargestellt werden. Hierzu wird unter dem Menüpunkt *Suchen / Anlegen* die Domäne und ein Pseudonym gewählt. Die entsprechende Hierarchie wird in einem Fenster dargestellt. Dabei sind die angewählte Domäne und das Pseudonym farblich hervorgehoben. In **Abbildung 5-11** ist die entsprechende Oberfläche dargestellt.

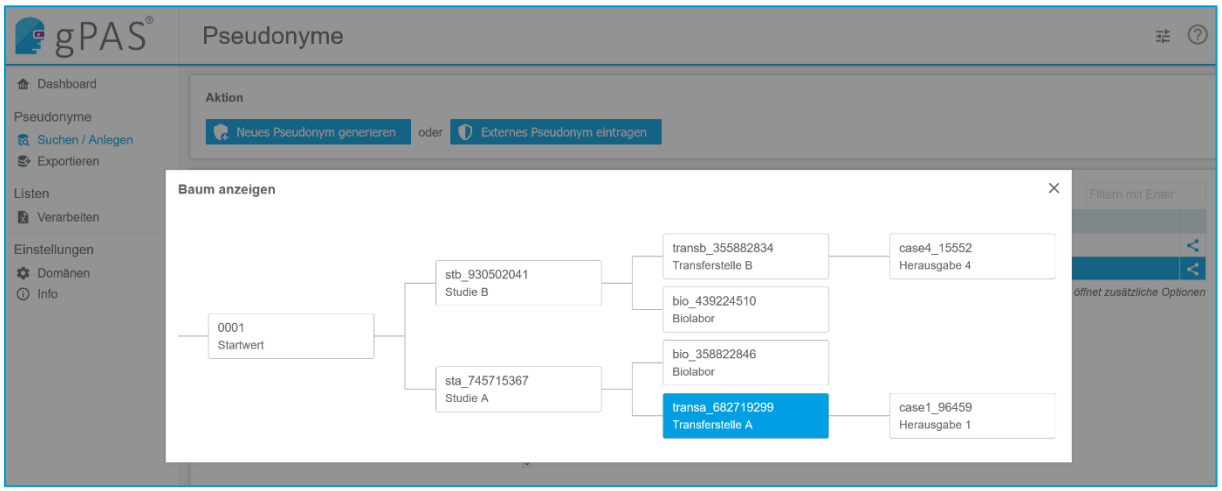



**Abbildung 5-11:** Oberfläche beim Anzeigen der Pseudonym-Hierarchie für ein selektiertes Pseudonym (blau hinterlegt).

Es werden neben dem gewählten Pseudonym auch der Originalwert und alle zugeordneten Pseudonyme dargestellt.

## 5.9 Anwendungsfall 9: Listenverarbeitung

Es ist möglich, eine Liste von Eingabewerten zu pseudonymisieren, zu depseudonymisieren (ermitteln des Originalwerts), zu anonymisieren oder zu löschen. Hierzu kann unter dem Menüpunkt *Verarbeiten* eine CSV-Datei ausgewählt werden. In **Abbildung 5-12** ist die entsprechende Oberfläche abgebildet.

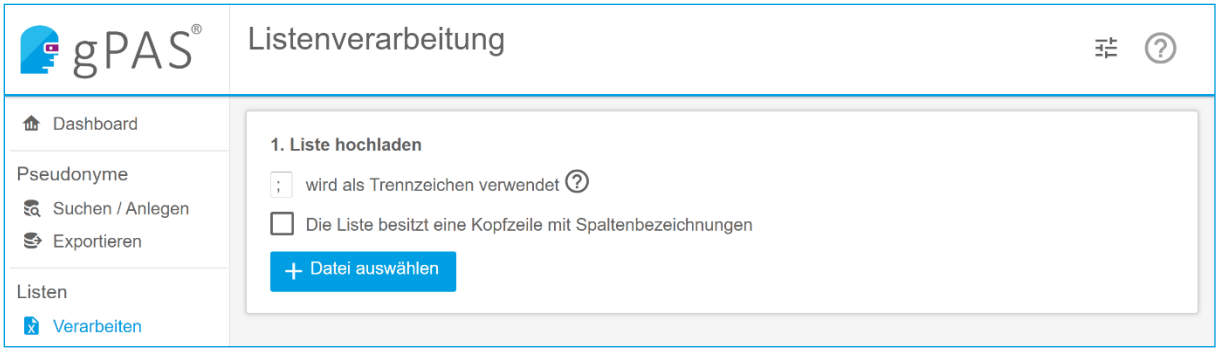



**Abbildung 5-12:** Oberfläche zum Verarbeiten von Listen.

Ist eine Überschrift enthalten, so kann dies mittels Anwählens der Checkbox *Die Liste besitzt eine Kopfzeile mit Spaltenbezeichnung* mitgeteilt werden. In diesem Fall wird die Kopfzeile/erste Zeile nicht mitverarbeitet. Wenn eine mehrspaltige Tabelle enthalten ist, erfolgt eine Separierung der Spalten standardmäßig mit einem Semikolon. Soll ein eigenes Trennzeichen verwendet werden, so kann dies entweder mittels `sep=X` (`X` steht für das verwendete Trennzeichen) in der ersten Zeile der CSV-Datei definiert werden oder in dem entsprechenden Feld das Trennzeichen definiert werden. Beim Hochladen erkennt der gPAS® automatisch die Kodierung der Datei. Diese kann bei Bedarf angepasst werden. Danach kann die Spalte gewählt werden, dessen Werte verarbeitet werden sollen. Die enthaltenen Werte werden dabei bereits in der Oberfläche als Vorschau angezeigt (vgl. **Abbildung 5-13**).

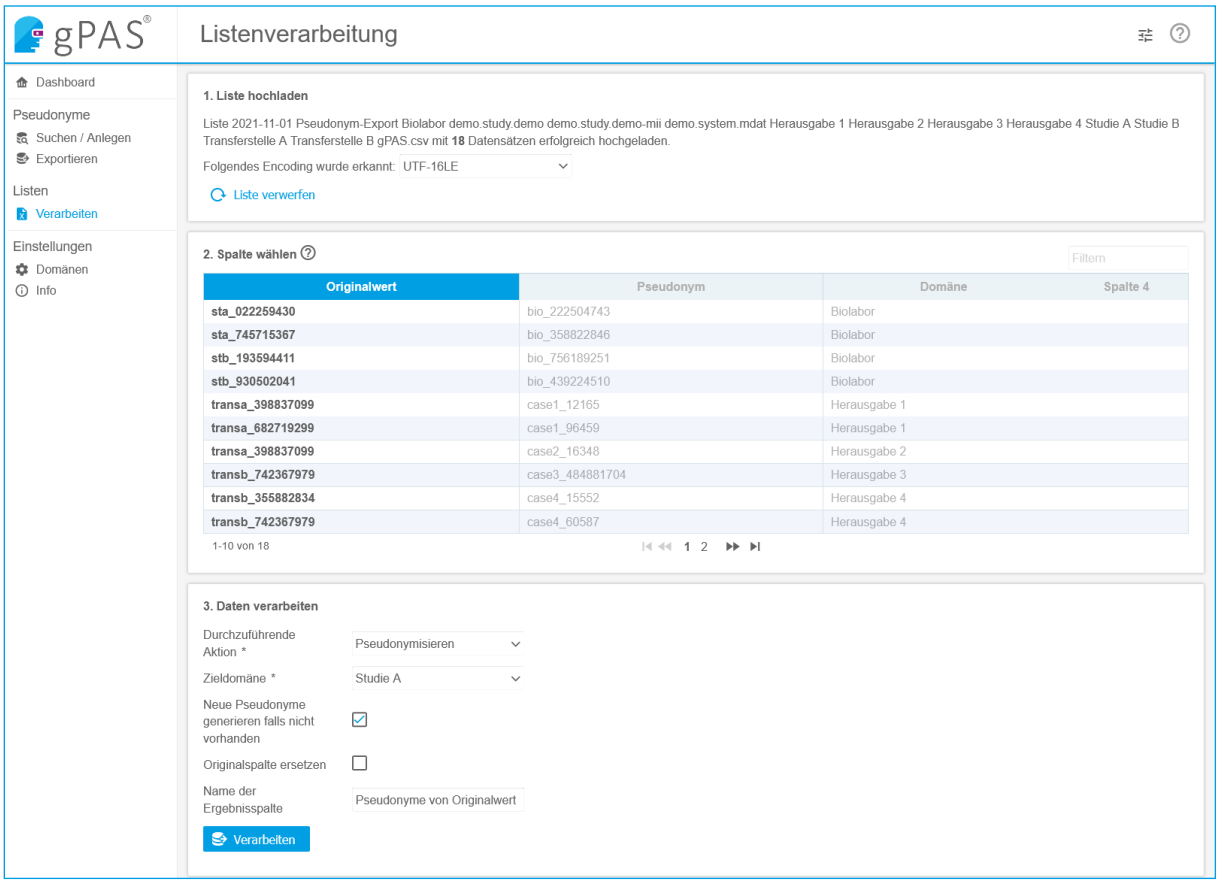



**Abbildung 5-13:** Wählen der Verarbeitungsoperation. Hier am Beispiel von *Pseudonymisieren*.

Es kann zwischen vier Verarbeitungsoperationen gewählt werden. Die entsprechenden Optionen sind in **Tabelle 5-4** aufgelistet.

**Tabelle 5-4:** Mögliche Verarbeitungsoperationen.

| Operation | Beschreibung |
|---|---|
| Pseudonymisieren | Für jedes Element der Liste ein neues Pseudonym erzeugen (sofern noch nicht bekannt). |
| Depseudonymisieren | Für jedes Pseudonym der Liste die Originalwerte ermitteln (sofern bekannt). |
| Anonymisieren | Jedes Pseudonym der Liste wird anonymisiert. |
| Löschen | Jedes Pseudonym der Liste wird aus dem Bestand entfernt. |

Die zu nutzende Pseudonym-Domäne muss angegeben werden.

Nach dem Wählen der Schaltfläche *Verarbeiten* wird das Ergebnis der Verarbeitung in die dargestellte Tabelle ergänzt. Hierfür kann die Ergebnisspalte benannt werden. Die aktualisierte Liste kann im Anschluss als CSV-Datei heruntergeladen werden.

## 5.10 Anwendungsfall 10: Dashboard für Statistiken

Die Anzahl der vorhandenen Pseudonyme, Anonyme und Pseudonym-Domänen im gPAS® können angezeigt werden und als CSV-Datei zum aktuellen Stand oder mit Verlaufsdaten exportiert werden. Unter dem Menüpunkt *Dashboard* können Tabellen und Diagramme eingesehen werden, welche die jeweiligen Daten aufbereitet darstellt. In **Abbildung 5-14** ist die Oberfläche der Statistik abgebildet.

Die gezeigten Statistiken werden asynchron, also nicht automatisch und nicht in Echtzeit, generiert. Die Aktualisierung kann jederzeit manuell angestoßen werden über die Schaltfläche „Aktualisieren“. Die dabei generierten Daten werden durch den gPAS erzeugt und in der Datenbank dokumentiert. Das Dashboard ersetzt damit die bis gPAS 1.11.x genutzten Kennzahlenprozeduren innerhalb der gPAS-Datenbank.

Die Statistik kann als CSV über die jeweiligen Schaltflächen heruntergeladen werden.

**ⓘ Unterstützung bei regelmäßiger Community-Kennzahlenerhebung.**

Das Dashboard liefert einen schnellen Überblick über Zahlen zu Pseudonymen und Domänen. Diese können als CSV-Datei exportiert und der Unabhängigen Treuhandstelle Greifswald per E-Mail übermittelt werden. Das unterstützt bei statistischen Auswertungen über die Gesamtzahl von Pseudonymen und Domänen in der Community. Vielen Dank fürs Mitmachen!

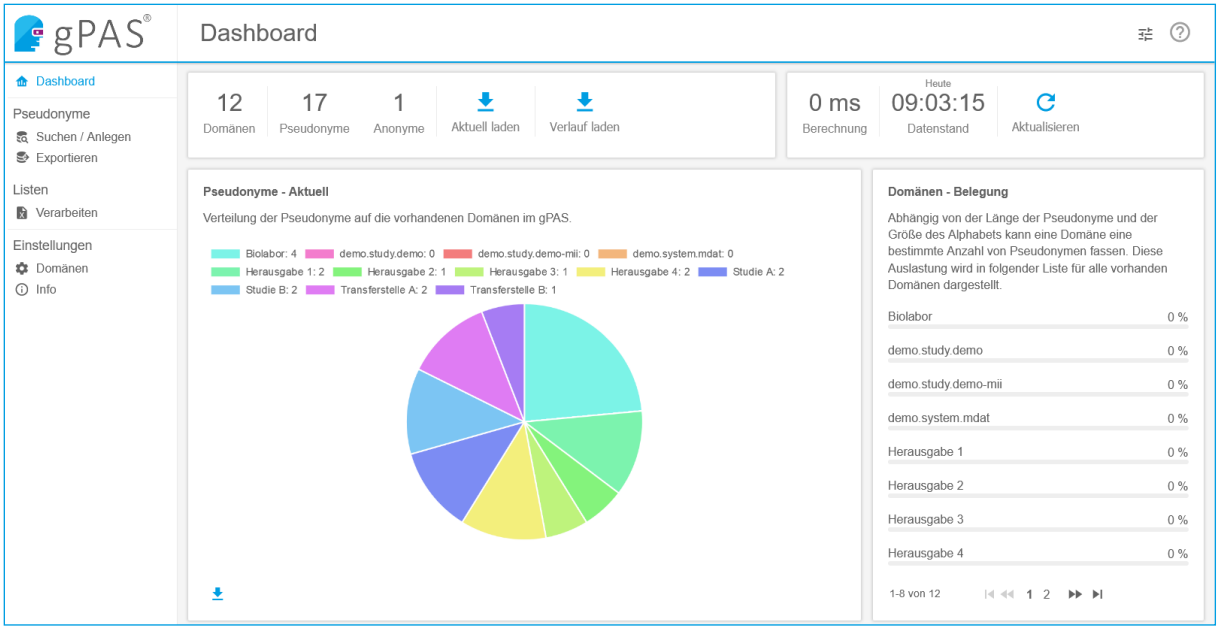



**Abbildung 5-14:** Oberfläche zum Einsehen von der Anzahl von Pseudonymen, Anonymen und Domänen. Die Daten sind in Diagrammen aufgeführt.

## 5.11 Anwendungsfall 11: Domäne bearbeiten oder löschen

Die Voraussetzung zum Bearbeiten oder Löschen einer Domäne ist, dass keine Pseudonyme in der entsprechenden Domäne hinterlegt sind. Um eine Domäne zu bearbeiten oder zu löschen, wird unter dem Menüpunkt *Domänen* die entsprechende Domäne ausgewählt und mit einem Rechtsklick das Kontextmenü aufgerufen (vgl. **Abbildung 5-15**).

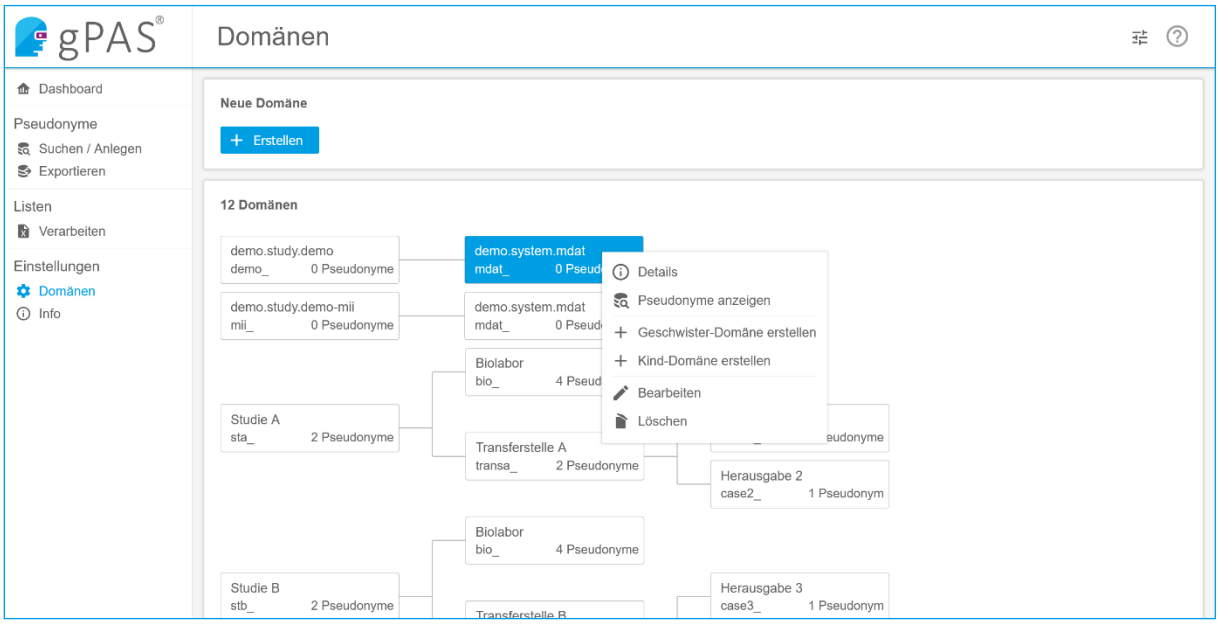



**Abbildung 5-15:** Kontextmenü mit den Schaltflächen zum Anzeigen der Domänendetails, zum Bearbeiten der Domäne und zum Löschen der Domäne. Hierüber können weitere Domänen erzeigt werden (s. **Anwendungsfall 1: Anlegen einer Domäne**).

Zum Bearbeiten wird die Schaltfläche *Bearbeiten* angewählt. Daraufhin können alle Einstellungen, außer der Namen, zur Domäne bearbeitet werden. Wenn die Schaltfläche *Löschen* gewählt wird, muss der Vorgang bestätigt werden und die Domäne wird unwiederbringlich gelöscht.

⚠ **Hinweis:** Wird der Name der Domäne geändert, betrifft dies nur den angezeigten Namen bzw. das Label. Wird die Domäne über die SOAP-Schnittstelle angesprochen, muss der ursprüngliche Name verwendet werden.

## 5.12 Anwendungsfall 12: Pseudonyme exportieren

Die Pseudonyme und dessen Originalwerte können exportiert werden. Hierbei kann gewählt werden, aus welchen Domänen die Pseudonyme exportiert werden sollen. Der Export wird über den Menüpunkt *Exportieren* aufgerufen. Die zu exportierenden Domänen können hierbei einzeln angewählt werden. Mittels eines Klicks auf die Schaltfläche „>“, werden die zu exportierenden Domänen in der rechten Liste gesammelt. Mit einem Klick auf die Schaltfläche *CSV herunterladen* wird der Export gestartet. Die resultierende Datei kann ohne weitere Anpassungen beispielsweise wieder importiert werden. In **Abbildung 5-16** ist die entsprechende Oberfläche dargestellt.

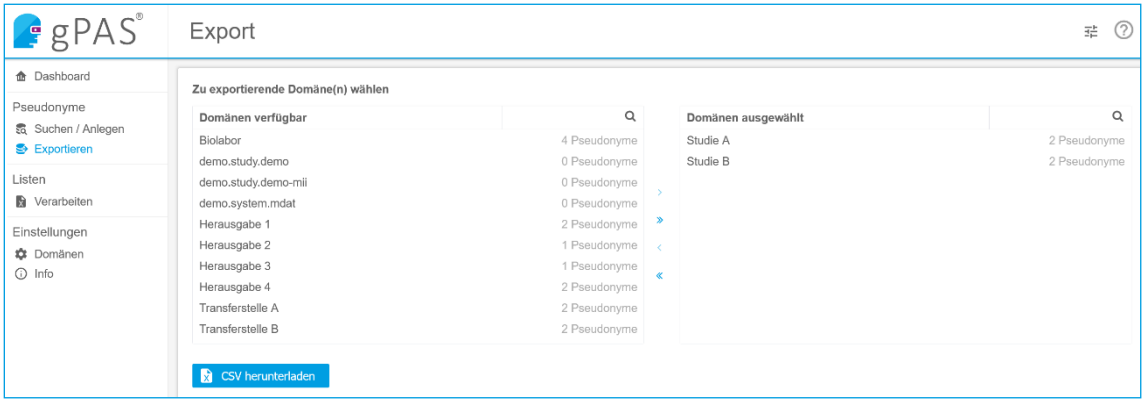



**Abbildung 5-16:** Oberfläche zum Exportieren beliebiger Domänen.

# 6 Logging

⚠ **Hinweis:** Details für die Anpassung der Logging-Konfiguration entnehmen Sie bitte der beigelegten Beschreibung docker-compose/README.md (Abschnitt Logging).

# 7 Versand von Notifications

Wie in der Architekturgrafik zu sehen (siehe **Abbildung 4-1**), ist der gPAS® seit Version 1.12.0 in der Lage Benachrichtigungen an externe Systeme zu versenden. Dies kann per `http`, `MQTT` oder `EJB` erfolgen. Die Versandmitteilungen werden in einer separaten Notification-Datenbank dokumentiert.

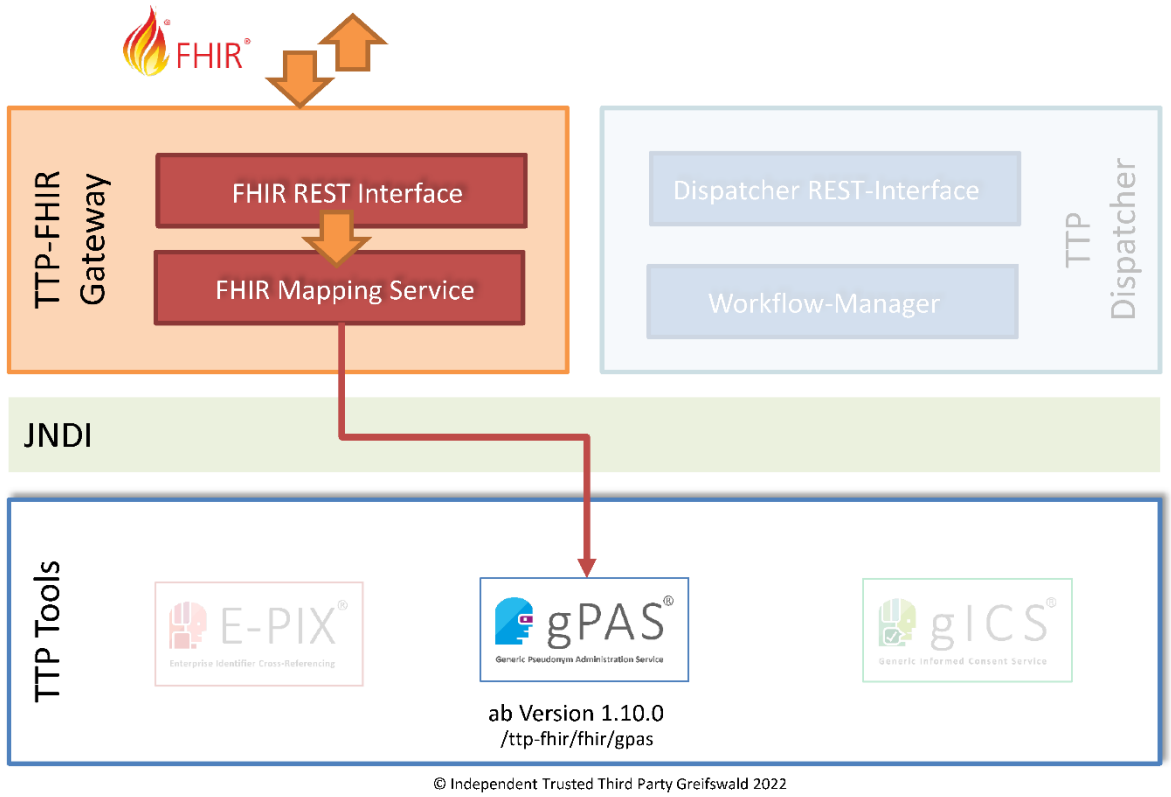
⚠ **Hinweis:** Details zum Umfang der Notification-Schnittstelle, zur Einrichtung, sowie weitere Erläuterungen sind separat unter https://www.ths-greifswald.de/ttp-tools/notifications dokumentiert.

# 8 FHIR-Unterstützung für gPAS® per TTP-FHIR Gateway

*„Fast Healthcare Interoperability Resources (kurz: FHIR®) ist ein von HL7 erarbeiteter Standard. Er unterstützt den Datenaustausch zwischen Softwaresystemen im Gesundheitswesen. FHIR beschreibt Datenformate und Elemente als sogenannte „Ressourcen" und bietet eine Schnittstelle an, um diese auszutauschen"*[8].

Um sowohl bestehende Anwenderprojekte als auch künftige Nutzer bei der Umsetzung FHIR-orientierter Infrastrukturen und Prozesse zu unterstützen, wird ab gPAS®-Version 1.10.0 ein zusätzliches Treuhandstellen-FHIR-Gateway (kurz: TTP-FHIR Gateway) als Mittler von FHIR-spezifischen Infrastrukturkomponenten und gPAS® bereitgestellt.

Für ausgewählte Funktionalitäten zur domänenspezifischen Generierung neuer Pseudonyme und zur Abfrage von Mappings von Pseudonym- und Originalwerten wurden in FHIR Funktionen umgesetzt und sind nach erfolgreichem Deployment direkt per REST nutzbar. Eine Liste der umgesetzten Funktionen ist unter https://ths-greifswald.de/gpas/fhir zu finden.




[8] https://de.wikipedia.org/wiki/Fast_Healthcare_Interoperability_Resources, Zugriff am 22.01.2021

⚠ **Hinweis:** Die Profilierung der erforderlichen Profile, Codesysteme und Operations erfolgte in Zusammenarbeit mit der Fa. gefyra[9].

Details zur Nutzung der Funktionen (Aufruf, Parameter, beispielhafte Antworten) können unter dem frei verfügbaren Implementation-Guide entnommen werden.

# 9 Authentifizierungs- und Autorisierung

Die bereitgestellte gPAS®-Version (ab 1.10.0) bietet unterschiedliche Umsetzungsoptionen der Authentifizierungs- und Autorisierung sowohl in der Docker- als auch in der Docker-Compose-Variante.

In der Standard-Ausgabe vom gPAS® ist keine Authentifizierung notwendig. Soll der gPAS® nur für bestimmte Nutzergruppen (Admin-Nutzer, Standard-Nutzer) zugänglich gemacht werden (vgl. Tabelle 9-1) oder das Anlegen von neuen Domänen beschränkt werden, stehen dafür zwei Authentifizierungsverfahren bereit. gRAS und KeyCloak, wobei es für KeyCloak zwei verschiedene Varianten gibt. *Die Verwendung von KeyCloak wird empfohlen*.

⚠ **Hinweis:** Die Auswahl der einzelnen Varianten erfolgt in der Docker-Compose Version innerhalb der docker-compose.yml. Details für die notwendige Anpassung der Docker-Konfiguration können der beigelegten Beschreibungen https://www.ths-greifswald.de/ttp-tools/keycloak sowie docker-compose/README.md.

⚠ **Hinweis:** Mit dem Herbstrelease 2022 können nun **alle THS-Schnittstellen (WEB-Oberfläche, FHIR-Gateway und SOAP-Webservices)** je Endpunkt und somit je Werkzeug (E-PIX, gICS, gPAS) mit KeyCloak-basierter (und damit OIDC-konformer) Authentifizierung abgesichert werden.

Die Konfiguration der Authentifizierung erfolgt in der Docker-Compose Version innerhalb der `ttp_gpas.env`. Eine detailliierte Beschreibung ist unter https://www.ths-greifswald.de/ttp-tools/keycloak verfügbar.

## 9.1 Übersicht Nutzerrollen und Rechte

**Tabelle 9-1:** Zugriffsrechte für Nutzer der Gruppe Admin und User in der Web-Oberfläche.

| Bereich/Seite | Zugang ohne Login | Zugang mit User-Rechten | Zugang mit Admin-Rechten |
|---|---|---|---|
| *Einstellungen: Info* | X | X | X |
| *Einstellungen*: Domänen | | | X |
| *Einstellungen*: Statistiken | | X | X |

[9] https://www.gefyra.de/, Zugriff am 2021-06-08

| | | |
|---|---|---|
| Listen: Importieren | | X |
| Pseudonyme: Suchen/Anlegen | X | X |
| Pseudonyme: Exportieren | | X |

## 9.2 Verwendung von KeyCloak

⚠ **Hinweis:** Details zur Vorbereitung des KeyCloak-Servers sind unter https://www.ths-greifswald.de/ttp-tools/keycloak beschrieben.

Die Client-seitige KeyCloak-Konfiguration kann sowohl per Config-Datei als auch per Environment-Variablen bei Start des Docker-Compose erfolgen. Details sind in **docker-compose/README.md** beschrieben.

Neben der Absicherung der Weboberfläche gibt es die Möglichkeit, die SOAP-Schnittstelle per KeyCloak abzusichern. Hierfür wird ähnlich wie bei der Weboberfläche in Zugriffsrechte für Admin und User unterschieden.

## 9.3 Verwendung von gRAS

⚠ **Hinweis:** Details zur Administration und Nutzung der gRAS-Authentifizierung sind unter folgendem Link https://www.ths-greifswald.de/ttp-tools/gras am Beispiel von gPAS® dokumentiert.

# 10 Empfehlungen zur Absicherung des Anwendungsservers

Der Zugriff auf relevante Anwendungs- und Datenbankserver des gPAS® sollte nur für autorisiertes Personal und über autorisierte Endgeräte möglich sein. Wir empfehlen die Umsetzung nachfolgender IT-Sicherheitsmaßnahmen:

- Betrieb der relevanten Server in separaten Netzwerkzonen (getrennt von Forschungs- und Versorgungsnetz)
- Verwendung von Firewalls und IP-Filtern
- Verwendung von KeyCloak
- Zugangsbeschränkung auf URL-Ebene mit Basic Authentication (z.B. mit NGINX oder Apache)

# 11 Nutzung der SOAP-Schnittstelle

Neben der grafischen Benutzerschnittstelle, stehen maschinenverständliche Web-Schnittstellen für den Regelbetrieb (Pseudonymisieren, Depseudonymisieren, Anonymisieren, Validieren, Strukturdarstellung) und für die Administration (Anlegen, Lesen und Löschen von Domänen) von gPAS® zur Verfügung. Diese können mit dem SOAP-Protokoll angesprochen werden. Beim laufenden Dienst

kann die Definition der SOAP-Schnittstelle mit den folgenden Pfaden, an der jeweiligen URL, abgerufen werden.

**Regelbetrieb**

```
http://example.org:8080/gpas/gpasService?wsdl
```

**Administration**

```
http://example.org:8080/gpas/DomainService?wsdl
```

**Die neueste Entwicklerdokumentation ist unter der folgenden URL zu finden.**

```
https://www.ths-greifswald.de/gpas/doc
```

## 11.1 Anlegen einer Domäne

Im Folgendem wird das Anlegen einer Domäne über die `DomainService`-Schnittstelle exemplarisch vorgestellt. In **Abbildung 11-1** ist die entsprechende Anfrage für das Anlegen einer Domäne dargestellt. Mittels des Parameters `name` wird ein eindeutiger Name der Domäne festgelegt. Der Prüfziffernalgorithmus wird mit dem Parameter `checkDigitClass` angegeben. Die verfügbaren Algorithmen können der **Tabelle 5-2** entnommen werden. Dabei muss vor den Namen des jeweiligen Algorithmus Paketpfad `org.emau.icmvc.ganimed.ttp.psn.generator` vorangestellt werden. Mit dem Parameter `alphabet` wird das verwendete Alphabet angegeben. Hierbei muss der Paketpfad `org.emau.icmvc.ganimed.ttp.psn.alphabets` vorangestellt werden. Die mitgelieferten Alphabete und die jeweils kompatiblen Prüfziffernalgorithmen, jeweils mit den kompletten Pfadangaben, sind in

**Tabelle** 11-1 aufgelistet. Daneben können auch eigene Alphabete angegeben werden. Mit dem Parameter `properties` werden die semikolonseparierten **Domain Properties** im Format SCHLÜSSEL=WERT übermittelt. Alle verfügbaren Eigenschaften sind in **Tabelle 5-3** aufgelistet. Falls die zu erstellende Domäne einer anderen Domäne untergeordnet werden soll (Kind-Domäne), so kann mittels des Parameters `parentDomainName` der Name der übergeordneten Domäne (Eltern-Domäne) angegeben werden.

**Tabelle 11-1:** Alphabete und die jeweils kompatiblen Prüfziffernalgorithmen.

| Alphabet | Kompatible Prüfziffernalgorithmen |
|---|---|
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.Hex* | |
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.Numbers* | org.emau.icmvc.ganimed.ttp.psn.generator.Verhoeff<br>org.emau.icmvc.ganimed.ttp.psn.generator.VerhoeffGumm<br>org.emau.icmvc.ganimed.ttp.psn.generator.Damm |
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.NumbersWithoutZero* | |
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.NumbersX* | org.emau.icmvc.ganimed.ttp.psn.generator.ReedSolomonLagrange |
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.Symbols31* | org.emau.icmvc.ganimed.ttp.psn.generator.ReedSolomonLagrange |
| *org.emau.icmvc.ganimed.ttp.psn.alphabets.Symbols32* | org.emau.icmvc.ganimed.ttp.psn.generator.HammingCode |

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/" xmlns:psn="http://psn.ttp.ganimed.icmvc.emau.org/">
   <soapenv:Header/>
   <soapenv:Body>
      <psn:addDomain>
         <domainDTO>
            <!--Optional:-->
            <name>Demo_Child_Domain</name>
            <comment>Zu Demo-Zwecken zur Anlage Kind-Domain</comment>
            <label>Demo_Domain_Child</label>
            <!--Optional:-->
            <checkDigitClass>org.emau.icmvc.ganimed.ttp.psn.generator.NoCheckDigits</checkDigitClass>
            <alphabet>org.emau.icmvc.ganimed.ttp.psn.alphabets.Symbol32</alphabet>
             <config>
               <!--Optional:-->
               <forceCache>DEFAULT</forceCache>
               <includePrefixInCheckDigitCalculation>false</includePrefixInCheckDigitCalculation>
               <includeSuffixInCheckDigitCalculation>false</includeSuffixInCheckDigitCalculation>
               <maxDetectedErrors>1</maxDetectedErrors>
               <psnLength>8</psnLength>
               <!--Optional:-->
               <psnPrefix>psn_</psnPrefix>
               <!--Optional:-->
            </config>
             <!--Zero or more repetitions:-->
            <parentDomainNames>Demo_Domain</parentDomainNames>
         </domainDTO>
      </psn:addDomain>
   </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 11-1:** XML-Repräsentation zum Anlegen einer Domäne über die SOAP-Schnittstelle.

Wurde die Domäne erfolgreich angelegt, antwortet der Dienst mit einer Antwort, wie sie in **Abbildung 11-2** dargestellt ist.

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
   <soap:Body>
      <ns2:addDomainResponse xmlns:ns2="http://psn.ttp.ganimed.icmvc.emau.org/"/>
   </soap:Body>
</soap:Envelope>
```

**Abbildung 11-2:** XML-Repräsentation der Rückgabe beim erfolgreichen Anlegen einer Domäne über die SOAP-Schnittstelle.

## 11.2 Anlegen von Pseudonymen

Im Folgendem wird das Anlegen von Pseudonymen mittels `gpasService`-Schnittstelle exemplarisch vorgestellt. Dabei kann mithilfe der Funktion `getOrCreatePseudonymFor` ein Pseudonym angelegt werden, wenn zum gegebenen Originalwert und Domäne noch kein Pseudonym vorhanden ist. Andernfalls wird das entsprechende Pseudonym zurückgeliefert. In **Abbildung 11-3** wird die eine exemplarische Anfrage dargestellt.

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:psn="http://psn.ttp.ganimed.icmvc.emau.org/">
   <soapenv:Header/>
   <soapenv:Body>
      <psn:getOrCreatePseudonymFor>
         <value>OriginalWert1</value>
         <domainName>Demo_Domain</domainName>
      </psn:getOrCreatePseudonymFor>
   </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 11-3:** XML-Repräsentation einer beispielhaften Anfrage zum Anlegen oder Abfragen eines Pseudonyms über die SOAP-Schnittstelle.

In **Abbildung 11-4** wird die entsprechende Antwort dargestellt. Darin enthalten ist das angelegte bzw. abgefragte Pseudonym.

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
   <soap:Body>
      <ns2:getOrCreatePseudonymForResponse
      xmlns:ns2="http://psn.ttp.ganimed.icmvc.emau.org/">
         <psn>1JTHCLAC</psn>
      </ns2:getOrCreatePseudonymForResponse>
   </soap:Body>
</soap:Envelope>
```

**Abbildung 11-4:** XML-Repräsentation der Rückgabe zu einer Anfrage zum Anlegen oder Abrufen eines Pseudonyms über die SOAP-Schnittstelle.

Zum Anlegen von Alt-Pseudonymen kann die Funktion `insertValuePseudonymPair` verwendet werden. Dabei wird unter Angabe des Originalwerts, des Pseudonyms und der Domäne der jeweilige Eintrag angelegt.

⚠ **Hinweis:** Insbesondere die hinzuzufügenden Pseudonyme müssen zwingend zur Konfiguration der verwendeten Pseudonym-Domäne passen und werden entsprechend validiert.

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
xmlns:psn="http://psn.ttp.ganimed.icmvc.emau.org/">
   <soapenv:Header/>
   <soapenv:Body>
      <psn:insertValuePseudonymPair>
         <value>OriginalWert2</value>
         <pseudonym>6JKLCMAM</pseudonym>
         <domainName>Demo_Domain</domainName>
      </psn:insertValuePseudonymPair>
   </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 11-5 XML-Repräsentation einer beispielhaften Anfrage zum Einfügen eines Originalwert-Pseudonym-Paares über die SOAP-Schnittstelle.**

Um mehrere Pseudonyme innerhalb einer Anfrage anzulegen (Batch-Verarbeitung), kann die Funktion `insertValuePseudonymPairs` verwendet werden.

## 11.3 Abfragen von (vorhandenen) Pseudonymen

Wenn ein bereits vorhandenes Pseudonym abgefragt werden soll, sozusagen ausschließlich lesend auf den gPAS® zugegriffen werden soll, kann die Funktion `getPseudonymFor` verwendet werden. Dabei werden der Originalwert und die Domäne angegeben. In **Abbildung 11-6** wird ein exemplarischer Abruf dargestellt.

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
      xmlns:psn="http://psn.ttp.ganimed.icmvc.emau.org/">
   <soapenv:Header/>
   <soapenv:Body>
      <psn:getPseudonymFor>
         <value>12AB34CD56</value>
         <domainName>KAS+</domainName>
      </psn:getPseudonymFor>
   </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 11-6:** XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Pseudonyms über die SOAP-Schnittstelle.

Die entsprechende Antwort wird in **Abbildung 11-7** dargestellt. Darin ist das entsprechende Pseudonym der Domäne zum angegebenen Originalwert enthalten.

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
   <soap:Body>
      <ns2:getPseudonymForResponse xmlns:ns2="http://psn.ttp.ganimed.icmvc.emau.org/">
         <psn>KAS_EVNT1XZJHJ</psn>
      </ns2:getPseudonymForResponse>
   </soap:Body>
</soap:Envelope>
```

**Abbildung 11-7:** XML-Repräsentation der Rückgabe zu einer Anfrage zum Abrufen eines Pseudonyms über die SOAP-Schnittstelle.

Mit der Funktion `getPseudonymForList` kann ein Abfragen mehrerer Pseudonyme mit nur einer Anfrage erreicht werden.

## 11.4 De-Pseudonymisieren (Abfragen von Originalwerten)

Ähnlich wie Pseudonyme, können Originalwerte mittels `gpasService`-Schnittstelle abgefragt werden. Hierzu kann mittels der Funktion `getValueFor` und unter Angabe des Pseudonyms und der Domäne der zugeordnete Originalwert abgefragt werden. In **Abbildung 11-8** ist eine Abfrage exemplarisch dargestellt.

```
<soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
      xmlns:psn="http://psn.ttp.ganimed.icmvc.emau.org/">
   <soapenv:Header/>
   <soapenv:Body>
      <psn:getValueFor>
         <psn>KAS_EVNT1XZJHJ</psn>
         <domainName>KAS+</domainName>
      </psn:getValueFor>
   </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 11-8:** XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Originalwerts über die SOAP-Schnittstelle.

In **Abbildung 11-9** wird die entsprechende Antwort dargestellt. Darin ist der Originalwert zum Pseudonym enthalten.

```
<soap:Envelope xmlns:soap="http://schemas.xmlsoap.org/soap/envelope/">
   <soap:Body>
      <ns2:getValueForResponse xmlns:ns2="http://psn.ttp.ganimed.icmvc.emau.org/">
         <value>12AB34CD56</value>
      </ns2:getValueForResponse>
   </soap:Body>
</soap:Envelope>
```

**Abbildung 11-9:** XML-Repräsentation der Rückgabe zu einer Anfrage zum Abrufen eines Originalwerts über die SOAP-Schnittstelle.

Um mehrere Originalwerte mit nur einer Anfrage abzufragen, kann die Funktion `getValueForList` verwendet werden.

# 12 Optimierungen

Wird entgegen der hier beschrieben Vorgehensweise selbst ein Applikationsserver und Datenbankserver aufgesetzt, so kann eine Performance-Steigerung des gPAS® durch diverse Optimierungen erzielt werden. In den von der Treuhandstelle Greifswald ausgelieferten Docker-Containern (WildFly und MySQL) sind diese bereits eingerichtet. Diese Optimierungen sind relevant, wenn mehr als 10 Mio. Datenbankeinträge erwartet werden.

## 12.1 Speicher für MySQL erhöhen

Standardmäßig ist im MySQL-Server eine `innodb_buffer_pool_size` von 128 MB eingestellt. Es wird empfohlen diese auf 2 GB zu erhöhen. Dies geschieht entweder direkt in der Datenbank oder bei

der Verwendung eines Docker-Containers als entsprechendes Kommando. Bei der Konfiguration dieses Wertes ist die offizielle MySQL-Dokumentation (https://dev.mysql.com/doc/refman/5.7/en/innodb-buffer-pool-resize.html) zu beachten. Die Anpassung dieses Wertes erfolgt unter Beachtung des verfügbaren Arbeitsspeichers.

## 12.2 Batch-Writing

Für jede Datenbankoperation (Insert, Update, Delete) wird standardmäßig separat auf die Datenbank zugegriffen. Zur Steigerung der Performance können die Anfragen jedoch zusammengefasst werden. Dies kann erreicht werden, indem in der `standalone.xml` des WildFly-Servers der Parameter `rewriteBatchedStatements=true` an die `jdbc-connection-url` angefügt wird.

## 12.3 Lange Zeiten zum Hochfahren des Applikationsservers

Wurden viele Millionen Pseudonyme angelegt und ein Neustart des Systems ist erforderlich, so kann das Hochfahren des Applikationsservers WildFly mehr Zeit in Anspruch nehmen, als der konfigurierte Timeout zulässt. Der Timeout wird standardmäßig nach 5 Minuten ausgelöst, sofern der WildFly bis dahin nicht hochgefahren ist. Es ist dann erforderlich, die Konfiguration des WildFly abzupassen. Hierzu wird in der `standalone.xml` des WildFly-Servers die Komponente `deployment-scanner` um das Attribut `deployment-timeout` ergänzt. Der Wert des Attributes gibt die Zeit in Sekunden an, ab wann ein Timeout ausgelöst wird. Im folgenden Beispiel wird der Timeout auf 15 Minuten (900 Sekunden) hochgesetzt.

```
<subsystem xmlns="urn:jboss:domain:deployment-scanner:2.0">
    <deployment-scanner […] scan-interval="5000"
    deployment-timeout="900" […]/>
</subsystem>
```

# 13 Upgrade

⚠ **Hinweis:** Manuelle Anpassungen bei einem Versionswechsel sind nur erforderlich, wenn der gPAS® in einem Applikationsserver betrieben wird und nicht die bereitgestellten Docker benutzt werden.

Beim Versionswechsel auf gPAS® 1.11.0 kann es beim Start und zyklisch im Betrieb zu Warnungen (`Failed to reinstate timer 'gpas.psn-ejb.StatisticManagerBean'`) im Serverlog kommen. Grund hier ist eine verschobene Timer-Funktion. Dies kann behoben werden, indem im Verzeichnis vom WildFly die unter `[WildFly-Verzeichnis]/standalone/data/timer-service-data` enthaltenen Dateien entfern werden. Danach kann der Dienst neu gestartet werden. Die neuen Dateien werden danach automatisch erzeugt und es kommt diesbezüglich zu keinen Warnungen mehr.

# 14 Publikationen und Vorträge

Bialke M*, Bahls T, Havemann C, Piegsa J, Weitmann K, Wegner T, et al.
**MOSAIC. A modular approach to data management in epidemiological studies. (Originalartikel)**
METHODS OF INFORMATION IN MEDICINE. 2015; 54(4):364-371.
http://dx.doi.org/10.3414/ME14-01-0133

Bialke M*, Penndorf P, Wegner T, Bahls T, Havemann C, Piegsa J, et al.
**A workflow-driven approach to integrate generic software modules in a Trusted Third Party (Originalartikel)**
Journal of Translational Medicine. 2015; 13(176).
http://www.translational-medicine.com/content/13/1/176

Bialke M*, Rau H*, Thamm O, Schuldt R, Penndorf P, Blumentritt A, Gött R, Piegsa J, Bahls T, Hoffmann W.
**Toolbox for Research, or How to facilitate a central data management in small-scale research projects. (Originalartikel)**
Journal of Translational Medicine. (open access) 2018; 16(16). IF: 4.197 (2017)
http://rdcu.be/FynH

# 15 Weiterführende Informationen

**Produktbroschüre gPAS®**
https://www.ths-greifswald.de/gpas/produktbrief

**gPAS® Service Spezifikation**
https://www.ths-greifswald.de/gpas/doc

**gPAS® Demo**
https://ths-greifswald.de/gpas/demo

**Offizielles gPAS® Docker-Image**
https://www.ths-greifswald.de/forscher/gpas/#download

**Git-Repository (Stand MOSIAC-Projekt, aktuelle Version siehe Offizielles Docker-Image)**
https://github.com/mosaic-hgw/gpas

**Docker Installation**
https://docs.docker.com/install/

**Docker-Compose Installation**
https://docs.docker.com/compose/install/

**Docker Cheat Sheet**
https://www.docker.com/sites/default/files/d8/2019-09/docker-cheat-sheet.pdf

**Docker und Docker-Compose Cheat Sheet**
https://dev-eole.ac-dijon.fr/doc/cheatsheets/docker.html