

# Anwenderhandbuch



Generic Pseudonym Administration Service

Version 1.0.2 vom 15.07.2019

*Herausgeber:*

Universitätsmedizin Greifswald K.d.ö.R.

*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](mailto:christopher.hampf@uni-greifswald.de)

## Versionierung

<b>Version</b>	<b>Datum</b>	<b>Bearbeitungsart / Betroffene Abschnitte</b>	<b>Bearbeiter</b>
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

## Inhalt

<b>Anwenderhandbuch</b> .....	<b>1</b>
<b>Versionierung</b> .....	<b>2</b>
<b>Inhalt</b> .....	<b>3</b>
<b>Abbildungsverzeichnis</b> .....	<b>4</b>
<b>Tabellenverzeichnis</b> .....	<b>5</b>
<b>1 Hintergrund</b> .....	<b>6</b>
<b>2 Begriffsbestimmungen</b> .....	<b>6</b>
<b>3 Funktionalitäten</b> .....	<b>7</b>
3.1 Was leistet der Dienst.....	7
3.2 Was leistet der Dienst nicht.....	8
<b>4 Installation</b> .....	<b>8</b>
4.1 Systemanforderungen.....	8
4.2 Download und Starten des Dienstes.....	8
<b>5 Die grafische Benutzeroberfläche des gPAS</b> .....	<b>10</b>
5.1 Anwendungsfall 1: Generieren von Pseudonymen.....	11
5.2 Anwendungsfall 2: Originalwerte und Pseudonyme suchen.....	14
5.3 Anwendungsfall 3: De-Pseudonymisierung.....	15
5.4 Anwendungsfall 4: Technische Anonymisierung.....	15
5.5 Anwendungsfall 5: Löschen von Pseudonymen.....	17
5.6 Anwendungsfall 6: Integration von Alt-Pseudonymen.....	17
5.7 Anwendungsfall 7: Anzeige von Pseudonym-Hierarchien.....	18
5.8 Anwendungsfall 8: Listenverarbeitung.....	18
5.9 Anwendungsfall 9: Statistik ermitteln.....	19
5.10 Anwendungsfall 10: Domäne bearbeiten oder löschen.....	20
<b>6 Nutzung der SOAP-Schnittstelle</b> .....	<b>20</b>
6.1 Anlegen von Pseudonymen.....	21
6.2 Abfragen von Pseudonymen.....	21
6.3 Abfragen von Originalwerten.....	22
<b>7 Publikationen und Vorträge</b> .....	<b>23</b>
<b>8 Weiterführende Informationen</b> .....	<b>24</b>

## Abbildungsverzeichnis

Abbildung 1-1: Zuweisung und Verwaltung von projektspezifischen Pseudonymen zu einem Personenidentifikator mittels gPAS. ....	6
Abbildung 4-1: Aktuelle Architektur des gPAS. ....	9
Abbildung 5-1: Oberfläche zum Anzeigen von allen angelegten Domänen.....	11
Abbildung 5-2: Oberfläche zum Anlegen einer neuen Domäne. ....	12
Abbildung 5-3: Oberfläche zum Anlegen eines neuen Pseudonyms. ....	14
Abbildung 5-4: Exemplarische Struktur bei mehreren Pseudonymen und Stufen für beispielsweise einen Studienteilnehmer.....	14
Abbildung 5-5: Oberfläche zum Suchen von Originalwerten oder Pseudonymen. ....	15
Abbildung 5-6: Mit einem Rechtsklick auf den Eintrag, kann das Kontextmenü aufgerufen werden..	16
Abbildung 5-7: Der unumkehrbare Vorgang der Anonymisierung muss nochmals bestätigt werden.	16
Abbildung 5-8: Anonymisierter Eintrag.....	17
Abbildung 5-9: Oberfläche beim Anzeigen der Pseudonymhierarchie.....	18
Abbildung 5-10: Oberfläche zum Verarbeiten von Listen. ....	18
Abbildung 5-11: Wählen der Verarbeitungsoperation.....	19
Abbildung 5-12: Oberfläche zum Einsehen von der Anzahl von Pseudonymen, Anonymen und Domänen. ....	20
Abbildung 5-13: Kontextmenü mit den Schaltflächen zum Einsetzen der Domänendetails, zum Bearbeiten der Domäne und zum Löschen der Domäne. ....	20
Abbildung 6-1: XML-Repräsentation einer beispielhaften Anfrage zum Anlegen oder Abfragen eines Pseudonyms über die SOAP-Schnittstelle des gPAS.....	21
Abbildung 6-2: XML-Repräsentation der Rückgabe zu einer Anfrage zum Anlegen oder Abrufen eines Pseudonyms über die SOAP-Schnittstelle.....	21
Abbildung 6-3: XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Pseudonyms über die SOAP-Schnittstelle des gPAS. ....	22
Abbildung 6-4: XML-Repräsentation der Rückgabe zu einer Anfrage zum Abrufen eines Pseudonyms über die SOAP-Schnittstelle. ....	22
Abbildung 6-5: XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Originalwerts über die SOAP-Schnittstelle des gPAS. ....	22
Abbildung 6-6: XML-Repräsentation der Rückgabe zu einer Anfrage zum Abrufen eines Originalwerts über die SOAP-Schnittstelle.....	23

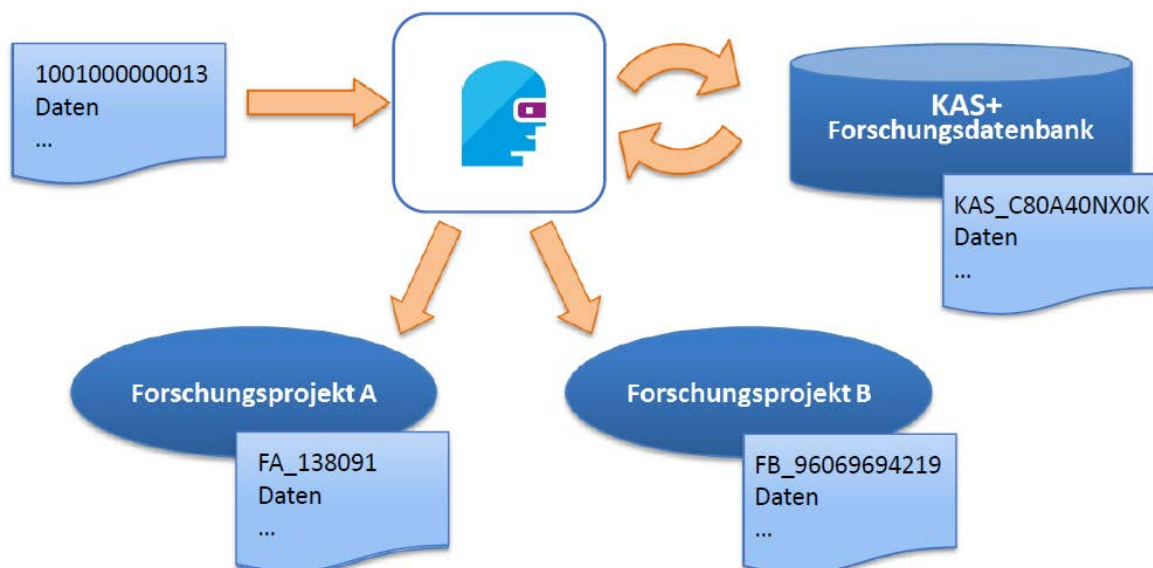
## Tabellenverzeichnis

Tabelle 5-1: Bereitgestellte Alphabete im gPAS.....	12
Tabelle 5-2: Mögliche Prüfziffern-Generatoren und dessen Bedingungen. ....	12
Tabelle 5-3: Mögliche Domain Properties.....	13
Tabelle 5-4: Mögliche Verarbeitungsoperationen. ....	19

## 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 lizenziert (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

### Domain (Domäne)

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.

## De-Pseudonymisieren

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 Pseudonymparametern: Prüfzifferalgorithmus, Länge, Alphabet
- Verwaltung von Pseudonymdomänen und Zweitpseudonymen
- Validierung von Pseudonymen
- Depseudonymisierung
- Darstellung von Pseudonymbäumen (-hierarchien)
- Import vorhandener Pseudonyme
- Löschen von temporären Pseudonymen (dabei werden sowohl Pseudonym als auch zugeordneter Originalwert gelöscht)
- Listenverarbeitung

## 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<sup>1</sup>

# 4 Installation

## 4.1 Systemanforderungen

### Technisch

- Windows oder Ubuntu Server (oder vergleichbar)<sup>2</sup>
- Mindestens 8 GB Arbeitsspeicher
- Mindestens 5 GB Festplattenspeicher
- Installierte aktuelle Version von *Docker*<sup>3</sup> und Docker Compose<sup>4</sup>
- Administrative Rechte
- Keine Nutzungsbeschränkungen auf die bereitgestellten Service- und Client-URLs

### 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

## 4.2 Download und Starten des Dienstes

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 (derzeit Docker Version 18.09.1 und Docker Compose Version 1.23.2) zu installieren.

---

<sup>1</sup> Nutzen Sie hierzu bitte ggf. den E-PIX: [thsgreifswald.de/e-pix](https://thsgreifswald.de/e-pix)

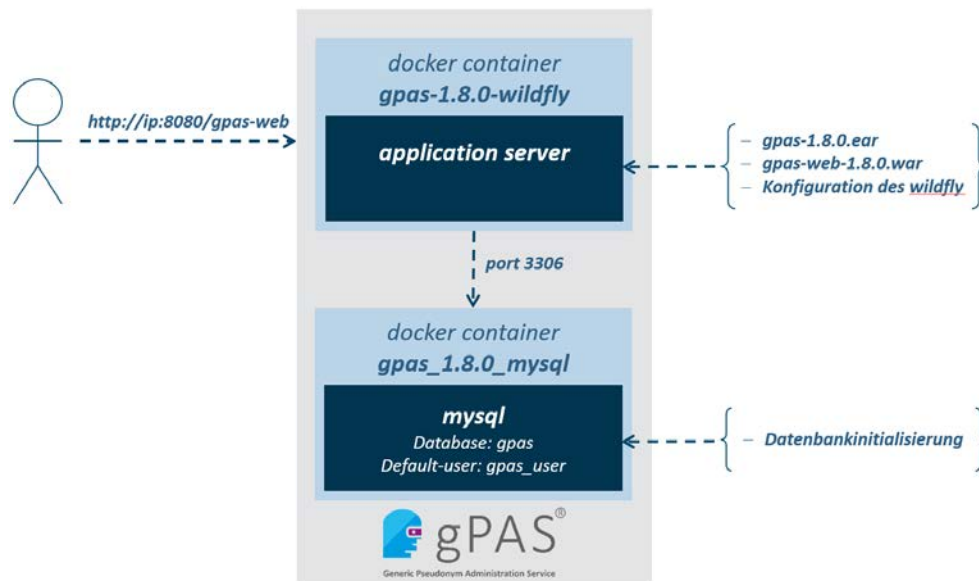
<sup>2</sup> Beim Betrieb unter Windows ist zu beachten, dass bei der Verwendung von Volumes und parallel betriebenen VPN-Clients Probleme auftreten können.

<sup>3</sup> Weitere Informationen unter <https://docs.docker.com/install/>

<sup>4</sup> Weitere Informationen unter <https://docs.docker.com/compose/install/>



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. Alle hierzu benötigten Ressourcen werden im öffentlichen Repository des MOSAIC-Projekts bereitgestellt<sup>5</sup>.



**Abbildung 4-1:** Aktuelle Architektur des gPAS.

Das entstehende Docker-System besteht aus getrennten Containern für die erforderliche Datenbankinstanz (MySQL) und den benötigten Anwendungsserver (Wildfly inkl. Datenbank-Konnektoren) zur Verfügbarmachung der Programmdateien. Beide Container kommunizieren intern über den MySQL-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.

Die Konfiguration erfolgt mittels Docker Compose. Der Installationsvorgang nimmt rund 3 Minuten in Anspruch und erfordern rund 1.4 GByte an Speicherplatz.

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

### Download der benötigten Dateien

```
sudo git clone https://github.com/mosaic-hgw/Dockerbank
```

### Vergabe Schreibrechten

```
sudo chmod -R 777 Dockerbank/gPAS/deployments
```

### Wechseln in das gPAS-Verzeichnis

```
cd Dockerbank/gPAS
```

<sup>5</sup> Weitere Informationen unter <https://github.com/mosaic-hgw/Dockerbank/tree/master/gPAS>  
Anwenderhandbuch

### Sollte der MySQL-Dienst auf der Maschine ausgeführt werden, dann stoppen des Dienstes

```
sudo service mysql stop
```

### Prüfen der Docker-Version (vorausgesetzt ist Version 1.13.1 oder höher)

```
sudo docker -v
```

### Prüfen der Docker Compose-Version (vorausgesetzt ist Version 1.8.0 oder höher)

```
sudo docker-compose -v
```

### Starten des gPAS mithilfe von Docker Compose

```
sudo docker-compose up
```

Damit werden die benötigten Komponenten heruntergeladen 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 7 Minuten dauern. Der erfolgreiche Start des Dienstes wird mit der folgenden Ausgabe abgeschlossen.

```
Wildfly 13.0.0.Final (Wildfly Core 5.0.0.Final) started in 63373ms -  
Started 824 of 1024 services (324 services are lazy, passive or on-demand)
```

## 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. Die daraus resultierenden Anwendungsfälle umfassen:

1. Generieren von Pseudonymen
2. Originalwerte und Pseudonyme suchen
3. De-Pseudonymisierung
4. Anonymisierung
5. Löschen von Pseudonymen
6. Integration von Alt-Pseudonymen
7. Anzeige von Pseudonym-Hierarchien
8. Verarbeitung von Listen
9. Statistik ermitteln
10. Domäne bearbeiten oder löschen

## 5.1 Anwendungsfall 1: Generieren von Pseudonymen

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

- ein Pseudonym ist innerhalb einer Pseudonymdomäne 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 Pseudonymparameter festlegen

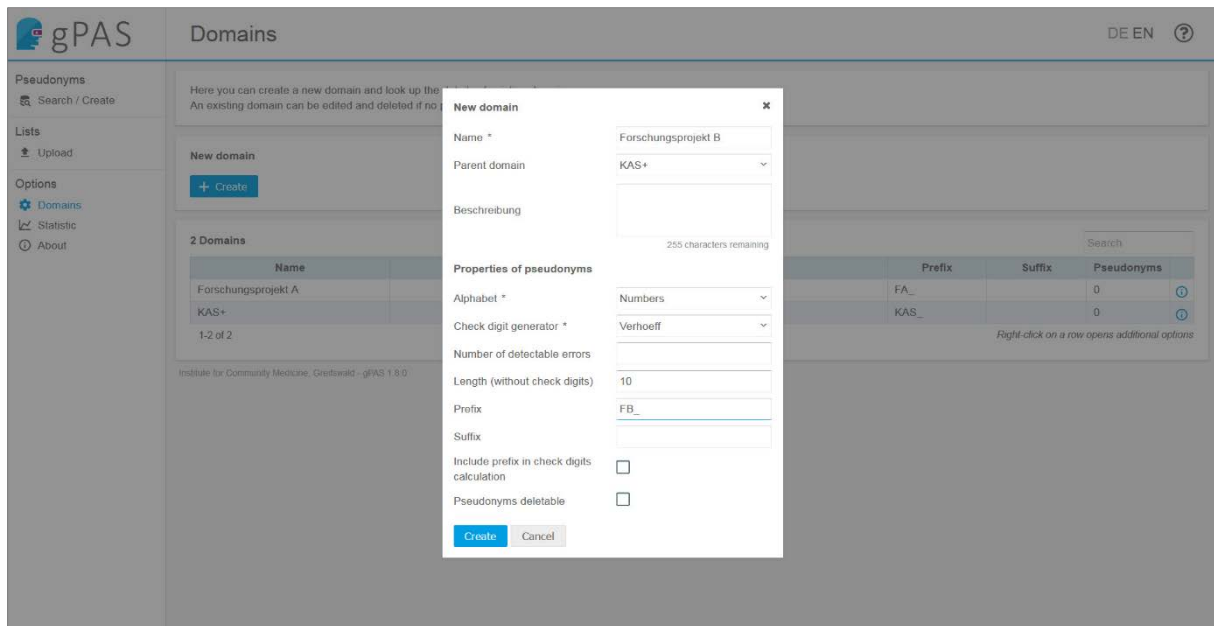
### Anlegen der Domäne

Unter dem Menüpunkt *Domänen* kann werden alle bereits angelegten Domäne aufgelistet (vgl. **Abbildung 5-1**). Um eine neue Domäne zu erzeugen, muss die Schaltfläche *Erzeugen* angewählt werden. Daraufhin öffnet sich ein Fenster mit den auszufüllenden Feldern.

Name	Parent domain	Beschreibung	Prefix	Suffix	Pseudonyms
Forschungsprojekt A		Datenelexport für Forschungsprojekt A	FA_		0
KAS+		Das KAS+Projekt	KAS_		0

**Abbildung 5-1:** Oberfläche zum Anzeigen von allen angelegten Domänen.

Beim Anlegen einer neuen Domäne ist der entsprechende **Name** festzulegen. Optional kann eine **Elterndomäne** (um Pseudonymhierarchien zu realisieren) und eine Beschreibung als **Kommentar** ergänzt werden. Zudem sind das zu nutzende **Alphabet** (vgl. **Tabelle 5-1**) und der jeweilige **Prüfziffern-Generator** (vgl. **Tabelle 5-2**) auszuwählen. 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 kommaseparierten, beliebigen Zeichen) definiert werden.

**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 Prüfzifferalgorithmen sind in **Tabelle 5-2** dargestellt. Alternativ kann die Option *Kein Prüfzeichen* genutzt werden.

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

Prüfzifferalgorithmus	Bedingung
HammingCode	Alphabet-Länge ist eine Primzahl
Verhoeff	Alphabet-Länge ist gleich 10

VerhoeffGumm	Alphabet-Länge ist gleich 10
Damm	Alphabet-Länge ist gleich 10
Reed-Solomon	Alphabet-Länge ist gleich Primzahlpotenz, wobei die maximale Anzahl der Prüfzeichen gleich der Alphabet-Länge ist

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

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

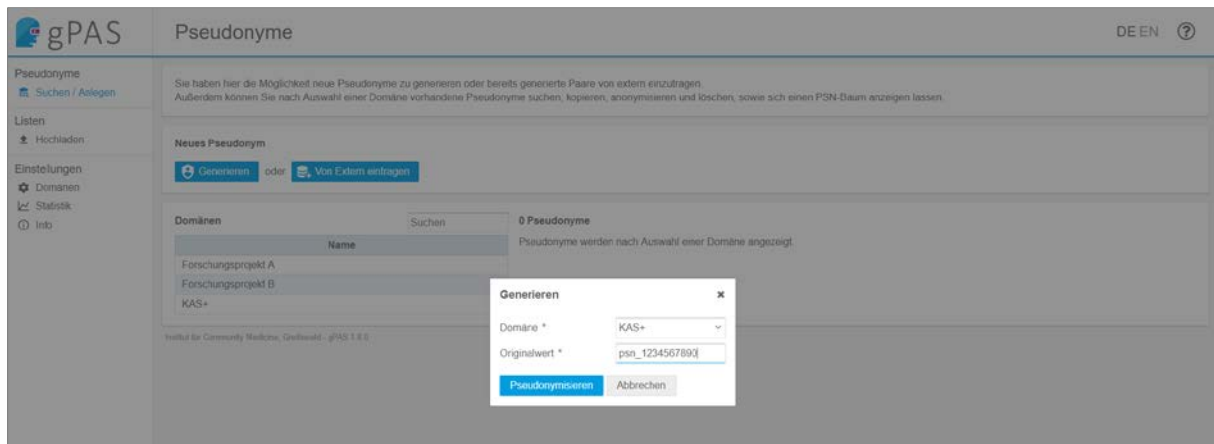
Alphabet	Bedeutung	Beispiel
MAX_DETECTED_ERRORS	Max. Anzahl fehlerhafter Zeichen, die auch sicher als fehlerhaft erkannt werden (nur bei Reed-Solomon)	2
PSN_LENGTH	Länge des erzeugten Pseudonyms (ohne Prüfzeichen)	10
PSN_PREFIX	Zeichenfolge, die als Präfix genutzt werden soll	abc_
PSN_SUFFIX	Zeichenfolge, die als Suffix genutzt werden soll	_xyz
INCLUDE_PREFIX_IN_CHECK_DIGIT_CALCULATION ↔	Einschließen des Präfix in die Berechnung der Prüfziffern	true, false
PSNS_DELETABLE	Pseudonyme innerhalb der Domäne sind löschar (z.B. zur Erstellung von temporären Pseudonymen, der Standard-Wert ist false.)	true, false

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 (`PSNS_DELETABLE=TRUE`).

### Generieren eines einzelnen Pseudonyms

Unter dem Menüpunkt *Suchen / Anlegen* können einzelne Pseudonyme angelegt werden. Hierzu muss zunächst die Schaltfläche *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.

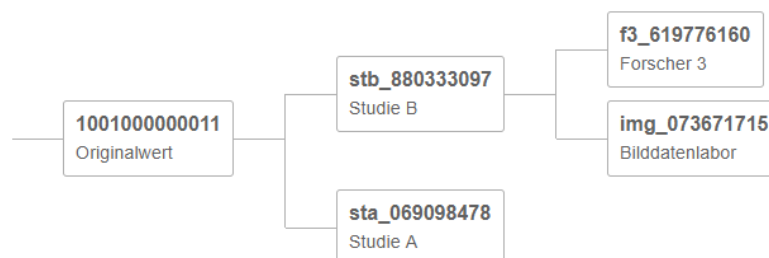


**Abbildung 5-3:** Oberfläche zum Anlegen eines neuen Pseudonyms.

### 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 genutzt.

In Fehler! Verweisquelle konnte nicht gefunden werden. wir 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 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.

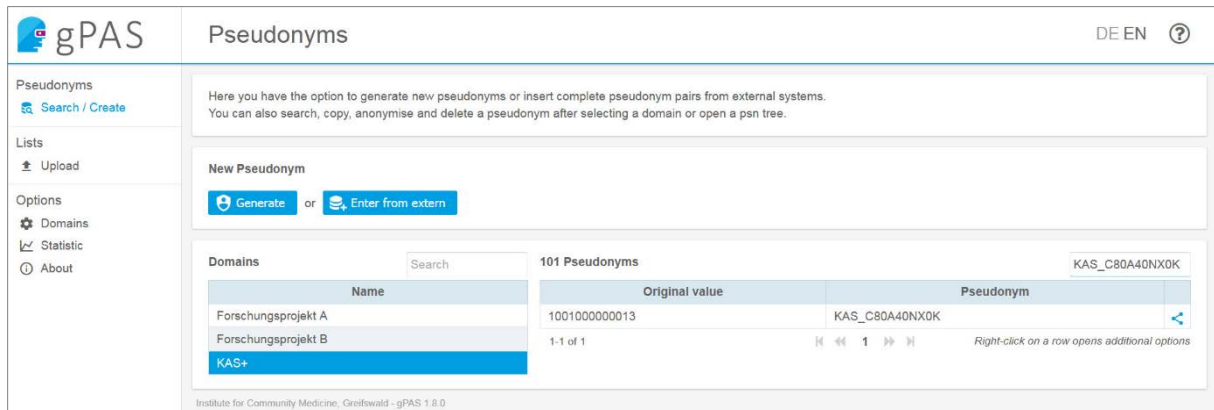


**Abbildung 5-4:** Exemplarische Struktur bei mehreren Pseudonymen und Stufen für beispielsweise einen Studienteilnehmer.

## 5.2 Anwendungsfall 2: 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 ein höherstufiges Pseudonym befindet. Wenn ein Pseudonym gesucht wird, dann wird entspricht die Domäne gewählt, in der sich das jeweilige

Pseudonym befindet. In das obere rechte Suchfeld (über der Pseudonym-Auflistung) wird die gesuchte Zeichenkette eingetragen. Dabei wird bereits beim Eintippen die Tabelle gefiltert. Es werden jedoch 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 5-5** ist die Oberfläche zum Suchen von Originalwerten und Pseudonymen dargestellt.



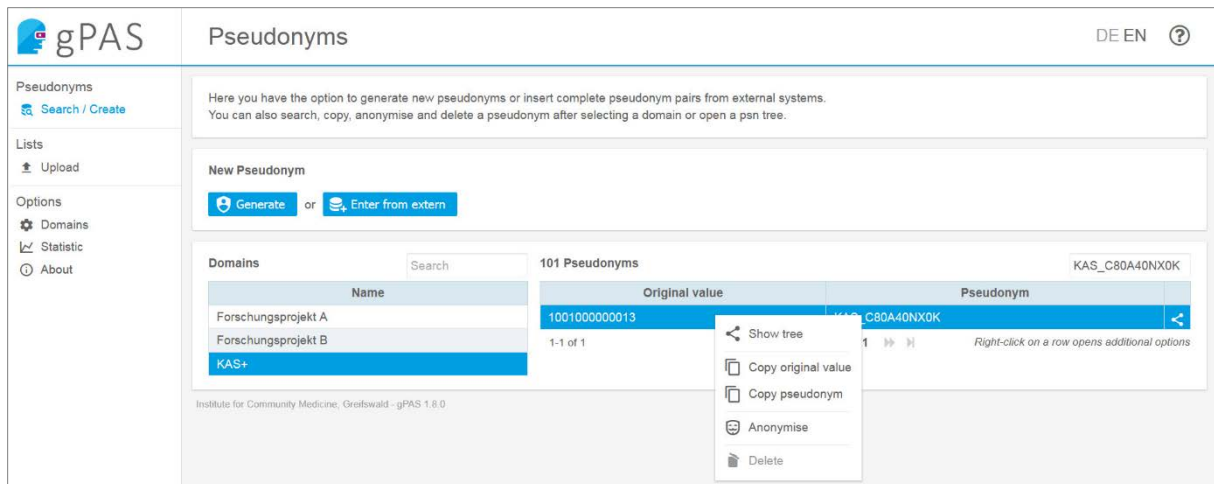
**Abbildung 5-5:** Oberfläche zum Suchen von Originalwerten oder Pseudonymen.

### 5.3 Anwendungsfall 3: De-Pseudonymisierung

Eine De-Pseudonymisierung entspricht der Suche eines Originalwerts zu einem bestimmten Pseudonym. Es wird demnach so vorgegangen, wie im **Abschnitt 5.2** beschrieben. Hierzu wird das vorhandene Pseudonym in der jeweiligen Domäne gesucht. Die Suche nach dem Pseudonym liefert den dazugehörigen Originalwert, respektive das Pseudonym geringerer Stufe.

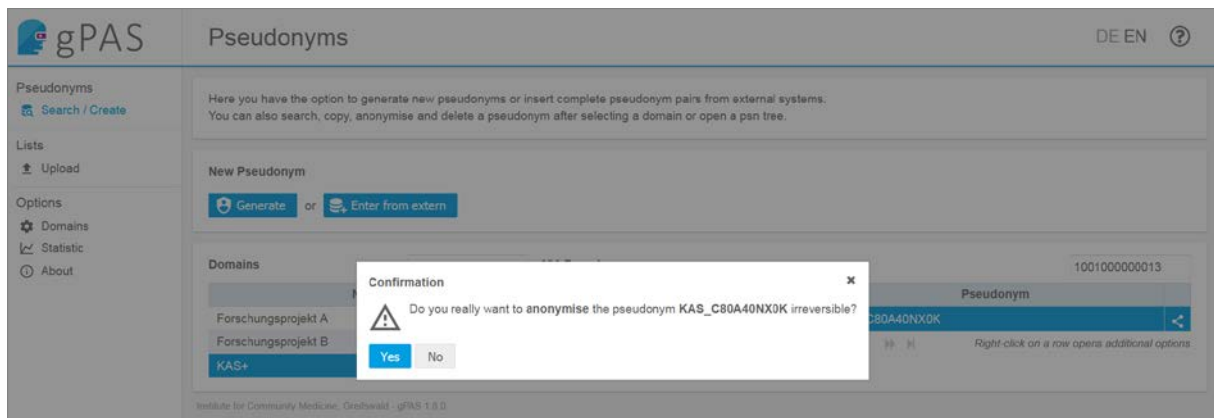
### 5.4 Anwendungsfall 4: Technische Anonymisierung

Bei der technischen Anonymisierung (im folgenden Anonymisierung genannt) 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 5.2**). Mit einem Rechtsklick auf den entsprechenden Eintrag wird das Kontextmenü geöffnet (siehe **Abbildung 5-6**).



**Abbildung 5-6:** Mit einem Rechtsklick auf den Eintrag, kann das Kontextmenü aufgerufen werden.

Mit dem Auswählen des Eintrags *Anonymisieren*, wird die Zuordnung aufgehoben. Bevor dieser unumkehrbare Vorgang ausgeführt wird, muss dies bestätigt werden (siehe **Abbildung 5-7**).



**Abbildung 5-7:** 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 5-8**).



The screenshot shows the gPAS Pseudonyms administration interface. At the top, there is a header with the gPAS logo, the title 'Pseudonyms', and language options 'DE EN' with a help icon. Below the header, there is a main content area with a sidebar on the left containing navigation options: 'Pseudonyms Search / Create', 'Lists Upload', 'Options Domains', 'Statistic', and 'About'. The main content area has a message: 'Here you have the option to generate new pseudonyms or insert complete pseudonym pairs from external systems. You can also search, copy, anonymise and delete a pseudonym after selecting a domain or open a psn tree.' Below this is an information box: 'Information: The original value 1001000000013 was deleted and the pseudonym KAS\_C80A40NX0K thereby successfully anonymised.' The main section is titled 'New Pseudonym' and has two buttons: 'Generate' and 'Enter from extern'. Below this is a table with two columns: 'Domains' and '101 Pseudonyms'. The 'Domains' table has a search bar and a list of domains: 'Forschungsprojekt A', 'Forschungsprojekt B', and 'KAS+'. The '101 Pseudonyms' table has a search bar and a list of pseudonyms with their original values. The first row is highlighted in blue, indicating it is selected. The table has a footer with '1-10 of 101' and navigation icons. A note at the bottom right says 'Right-click on a row opens additional options'.

Domains	Search	101 Pseudonyms	Search
	Name	Original value	Pseudonym
	Forschungsprojekt A	###_anonym_###_KTC6E98R5CUL_###_anonym_###	KAS_C80A40NX0K
	Forschungsprojekt B	1001063893033	KAS_264PKMWRXM
	<b>KAS+</b>	1001130368754	KAS_T0JPAQK4A4
		1001258786447	KAS_KYDU8P1MZY
		1001270131586	KAS_70NM4RADA6
		1001431391192	KAS_UN8MN36CXL
		1001462386428	KAS_T5AWZDCREN
		1001567939657	KAS_96ETCY2U44
		1001569905357	KAS_UJXZ43ADXG
		10016568399	KAS_EYZJ4336X4

Abbildung 5-8: Anonymisierter Eintrag.

## 5.5 Anwendungsfall 5: Löschen von Pseudonymen

Beim Löschen eines Pseudonyms, wird der Originalwert und das entsprechende Pseudonym unwiederbringlich aus der Domäne entfernt. Der Originalwert bleibt jedoch als Pseudonym niedriger Stufe bestehen. Eine De-Pseudonymisierung 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 5-6**). Mit dem Auswählen des Eintrags *Löschen*, wird der Eintrag gelöscht. Bevor dieser unumkehrbare Vorgang ausgeführt wird, muss dies bestätigt werden.

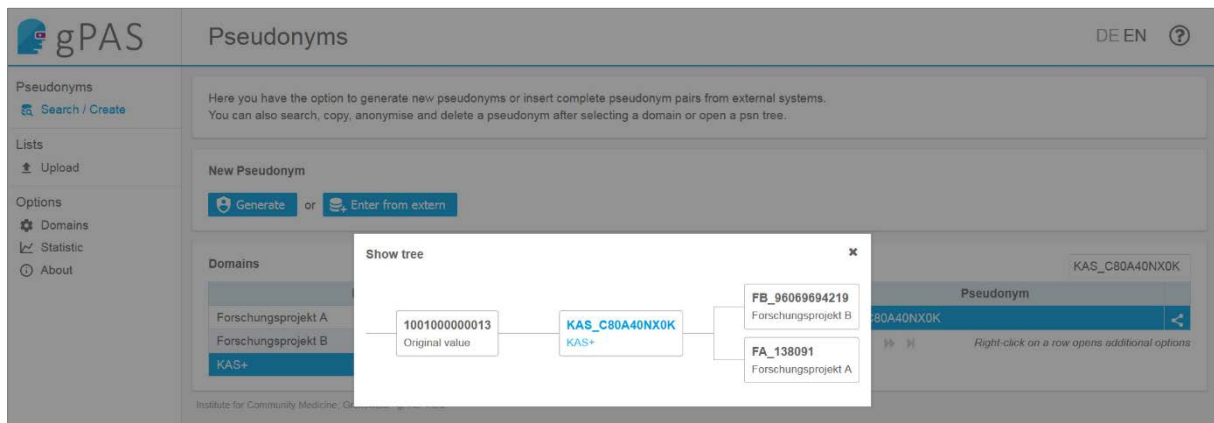
Es ist möglich, Pseudonyme zu löschen, die Pseudonyme höherer Stufe aufweisen. Wird dies getan, so bleibt das höherstufige Pseudonym samt Originalwert bestehen. Eine De-Pseudonymisierung ist dann jedoch nicht mehr möglich.

## 5.6 Anwendungsfall 6: 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.

## 5.7 Anwendungsfall 7: Anzeige von Pseudonym-Hierarchien

Wurden bei der Konfiguration der Pseudonym-Domäne eine Eltern-Domäne zur Angabe der Beziehung zweier Domänen (Eltern-Kind-Beziehung, 1:N möglich) angegeben, so kann der dadurch entstehende Pseudonymbaum 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 ist die angewählte Domäne und Pseudonym farblich hervorgehoben. In **Abbildung 5-9** ist die entsprechende Oberfläche dargestellt.

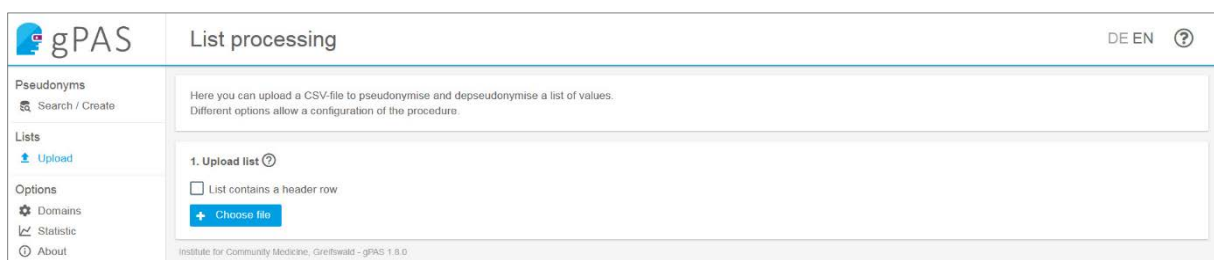


**Abbildung 5-9:** Oberfläche beim Anzeigen der Pseudonymhierarchie.

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

## 5.8 Anwendungsfall 8: Listenverarbeitung

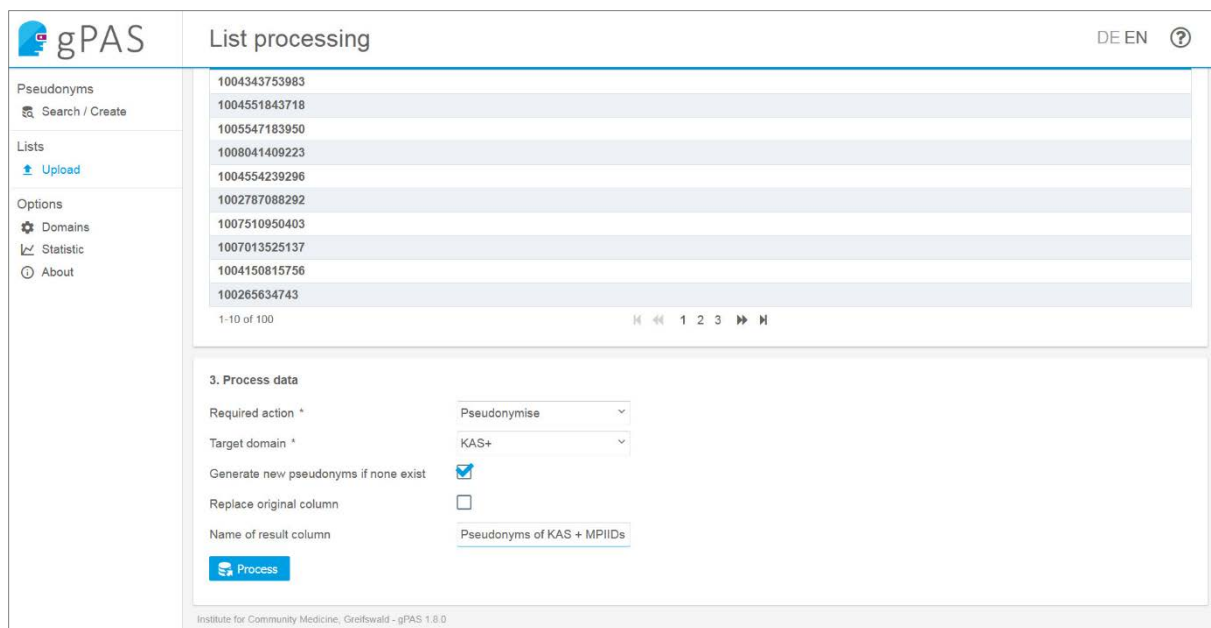
Es ist möglich, eine Liste von Eingabewerte zu pseudonymisieren oder zu de-pseudonymisieren (ermitteln des Originalwerts). Hierzu kann unter dem Menüpunkt *Hochladen* eine CSV-Datei ausgewählt werden. In **Abbildung 5-10** ist die entsprechende Oberfläche abgebildet.



**Abbildung 5-10:** Oberfläche zum Verarbeiten von Listen.

Ist eine Überschrift enthalten, so kann dies mittels Anwählen der Checkbox *Liste besitzt eine Kopfzeile* mitgeteilt werden. In diesem Fall wird die Kopfzeile nicht mitverarbeitet. Wenn eine mehrspaltige Tabelle enthalten ist, erfolgt eine Separierung der Spalten standardmäßig mit einem Komma. Soll ein eigenes Trennzeichen verwendet werden, so kann dies mittels `sep="X"` in der ersten Zeile der CSV-Datei definiert werden. Hierbei steht X für das verwendete Trennzeichen. Nachdem die Datei

importiert wurde, kann die Spalte gewählt werden, dessen Werte verarbeitet werden sollen. Die enthaltenen Werte werden dabei bereits in der Oberfläche als Vorschau angezeigt (vgl. **Abbildung 5-11**).



**Abbildung 5-11:** Wählen der Verarbeitungsoperation.

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

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

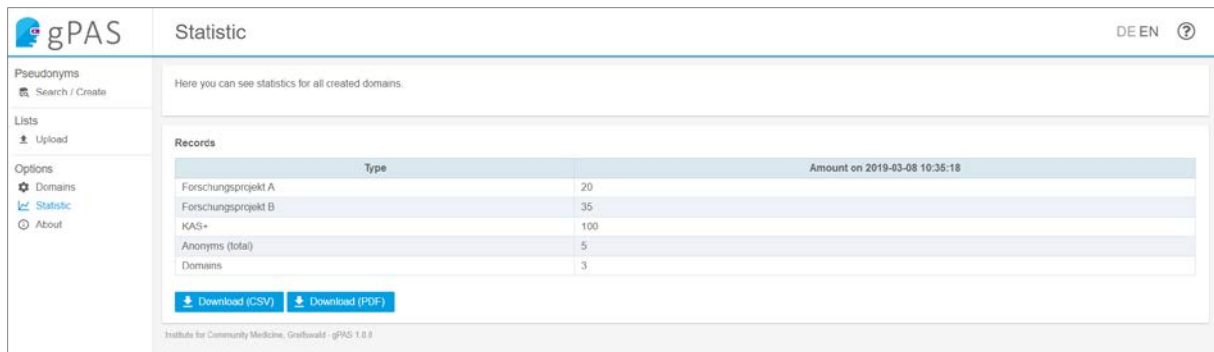
Operation	Beschreibung
Pseudonymisieren	Für jedes Element der Liste ein neues Pseudonym erzeugen (sofern noch nicht bekannt).
De-Pseudonymisieren	Für jedes Pseudonym der Liste die Originalwerte ermitteln (sofern bekannt).

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.

## 5.9 Anwendungsfall 9: Statistik ermitteln

Die Anzahl der vorhandenen Pseudonyme, Anonyme und Pseudonymdomänen im gPAS können angezeigt werden und als Ergebnis im CSV- oder PDF-Format exportiert werden. Hierzu kann unter dem Menüpunkt *Statistik* eine Tabelle eingesehen werden, welche die jeweiligen Daten auflistet. In **Abbildung 5-12** ist die Oberfläche der Statistik abgebildet.



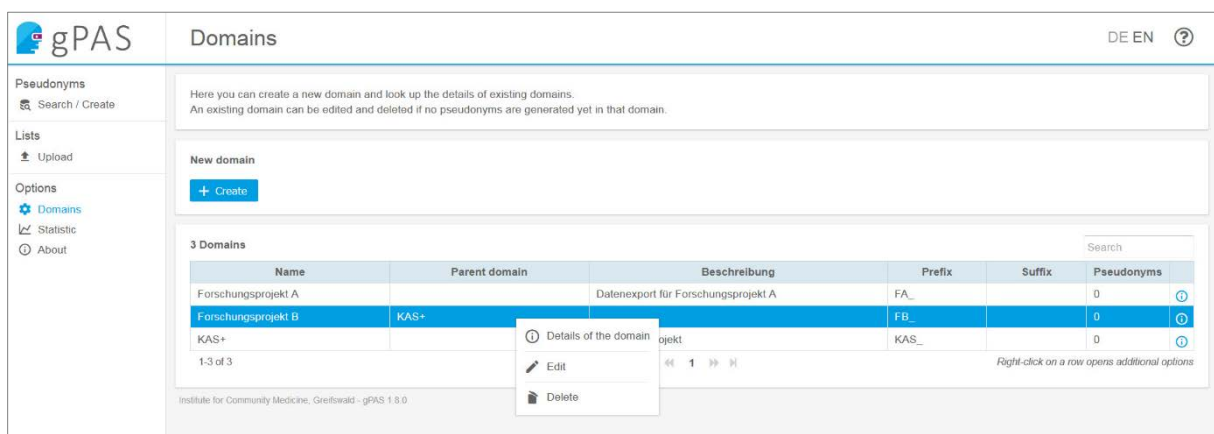
The screenshot shows the 'Statistic' page in gPAS. It features a sidebar with navigation options: Pseudonyms, Lists, and Options (Domains, Statistic, About). The main content area displays a table of records for domains created on 2019-03-08 10:35:18. Below the table are buttons for 'Download (CSV)' and 'Download (PDF)'. The footer indicates the software is from the Institute for Community Medicine, Greifswald, gPAS 1.8.0.

Type	Amount on 2019-03-08 10:35:18
Forschungsprojekt A	20
Forschungsprojekt B	35
KAS+	100
Anonyms (total)	5
Domains	3

**Abbildung 5-12:** Oberfläche zum Einsehen von der Anzahl von Pseudonymen, Anonymen und Domänen.

## 5.10 Anwendungsfall 10: Domäne bearbeiten oder löschen

Die Voraussetzung zum Bearbeiten oder Löschen einer Domäne ist, dass keine Pseudonyme bei 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 5-13).



The screenshot shows the 'Domains' page in gPAS. It includes a sidebar with navigation options: Pseudonyms, Lists, and Options (Domains, Statistic, About). The main content area features a 'New domain' section with a '+ Create' button and a table of existing domains. A context menu is open over the 'KAS+' row, showing options for 'Details of the domain', 'Edit', and 'Delete'. The footer indicates the software is from the Institute for Community Medicine, Greifswald, gPAS 1.8.0.

Name	Parent domain	Beschreibung	Prefix	Suffix	Pseudonyms
Forschungsprojekt A		Datenexport für Forschungsprojekt A	FA_		0
Forschungsprojekt B	KAS+		FB_		0
KAS+			KAS_		0

**Abbildung 5-13:** Kontextmenü mit den Schaltflächen zum Einsetzen der Domänendetails, zum Bearbeiten der Domäne und zum Löschen der 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.

## 6 Nutzung der SOAP-Schnittstelle

Neben der grafischen Benutzerschnittstelle, steht eine maschinenverständliche Web-Schnittstelle zur Verfügung. Diese kann mit dem SOAP-Protokoll angesprochen werden. Beim laufenden Dienst kann die Definition der SOAP-Schnittstelle mit dem folgenden Pfad, an der jeweiligen URL, abgerufen werden.

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

Die dazugehörige Entwicklerdokumentation ist unter der folgenden URL zu finden.

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

## 6.1 Anlegen von Pseudonymen

Im Folgendem wird das Anlegen von Pseudonymen 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 6-1** 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>12AB34CD56</value>
      <domainName>KAS+</domainName>
    </psn:getOrCreatePseudonymFor>
  </soapenv:Body>
</soapenv:Envelope>
```

**Abbildung 6-1:** XML-Repräsentation einer beispielhaften Anfrage zum Anlegen oder Abfragen eines Pseudonyms über die SOAP-Schnittstelle des gPAS.

In **Abbildung 6-2** 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>KAS_EVNT1XZJHJ</psn>
    </ns2:getOrCreatePseudonymForResponse>
  </soap:Body>
</soap:Envelope>
```

**Abbildung 6-2:** 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. Um mehrere Pseudonyme innerhalb einer Anfrage anzulegen, kann die Funktion `insertValuePseudonymPairs` verwendet werden.

## 6.2 Abfragen von Pseudonymen

Wie in **Abschnitt 6.1** gezeigt, kann mittels der Funktion `getOrCreatePseudonymFor` ein Pseudonym angelegt oder wenn dieses bereits vorhanden ist, abgefragt werden. Wenn ein Pseudonym abgefragt werden soll, ohne ein etwaiges Anlegen von diesem, kann die Funktion

getPseudonymFor verwendet werden. Dabei wird der Originalwert und die Domäne angegeben. In **Abbildung 6-3** 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 6-3:** XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Pseudonyms über die SOAP-Schnittstelle des gPAS.

Die entsprechende Antwort wird in **Abbildung 6-4** 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 6-4:** 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.

### 6.3 Abfragen von Originalwerten

Ähnlich wie Pseudonyme, können Originalwerte abgefragt werden. Hierzu kann mittels der Funktion getValueFor und unter Angabe des Pseudonyms und der Domäne der zugeordnete Originalwert abgefragt werden. In **Abbildung 6-5** 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 6-5:** XML-Repräsentation einer beispielhaften Anfrage zum Abfragen eines Originalwerts über die SOAP-Schnittstelle des gPAS.

In **Abbildung 6-6** 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 6-6:** 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.

## 7 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>

## 8 Weiterführende Informationen

### Produktbroschüre gPAS

<https://www.ths-greifswald.de/wp-content/uploads/2019/01/gPAS-Produktbrief.pdf>

### gPAS Service Spezifikation

<https://www.ths-greifswald.de/gpas/doc>

### gPAS Demo

<https://ths-greifswald.de/gpas/demo>

### Git-Repository

<https://github.com/mosaic-hgw/Dockerbank/tree/master/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/Docker\\_CheatSheet\\_08.09.2016\\_0.pdf](https://www.docker.com/sites/default/files/Docker_CheatSheet_08.09.2016_0.pdf)

### Docker und Docker Compose Cheat Sheet

<https://dev-eole.ac-dijon.fr/doc/cheatsheets/docker.html>