

# Leitlinie zur Installation und Konfiguration des E-PIX



## Erstellt von:

Lizon Fiedler-Lacombe,  
Pia Naumann

## Reviewed von:

Dr. Anika Kästner,  
Dr. Martin Bialke,  
Christopher Hampf

DATUM: 07.02.2024

VERSION 2.0

|         |  |    |
|---------|--|----|
| 1.      | Einleitung .....   | 3  |
| 2.      | Installation des E-PIX unter „Docker für Windows“ .....        | 4  |
| 3.      | Konfiguration des E-PIX via Oberfläche .....                   | 5  |
| 3.1     | Anlegen der Domänen und Quellen .....                          | 5  |
| 3.2     | Datenquelle anlegen (Sichere Quelle) .....                     | 6  |
| 3.3     | Domäne anlegen (Matching-Domäne) .....                         | 7  |
| 3.3.1   | Domäne manuell anlegen .....                                   | 7  |
| 3.3.1.1 | Anlegen einer Domäne: Reiter Einstellungen .....               | 7  |
| 3.3.1.2 | Anlegen einer Domäne: Reiter Personenfelder .....              | 8  |
| 3.3.1.3 | Anlegen einer Domäne: Reiter Vorverarbeitung .....             | 9  |
| 3.3.1.4 | Anlegen einer Domäne: Reiter Matching .....                    | 10 |
| 3.3.1.5 | Anlegen einer Domäne: Reiter Privatsphäre .....                | 12 |
| 3.3.2   | Domäne anlegen via XML-Datei .....                             | 14 |
| 4.      | Umwandlung der IDAT zu Bloomfiltern .....                      | 19 |
| 5.      | Exportieren der IDAT und Bloomfilter via Export-Funktion ..... | 21 |

## 1. Einleitung

Diese Leitlinie zur Installation und Konfiguration des E-PIX wurde im Rahmen des DigiNet-Projektes erstellt und stellt einen von mehreren möglichen Use Cases zur Nutzung des Dienstes dar. Details zum DigiNet-Projekt finden Sie bei Interesse unter:

<https://www.ths-greifswald.de/projekte/diginet/>

Der projektspezifische Use Case von DigiNet ist, dass Daten aus unterschiedlichen Datenquellen (Krebsregister und Krankenkassen) anhand der Krankenversicherernummern zusammengeführt werden sollen, ohne dass die Krankenversicherernummern von den entsprechenden Institutionen herausgegeben werden dürfen. Für diesen Zweck werden Bloomfilter erzeugt.

Die folgenden graphischen Darstellungen sowie die Konfiguration des E-PIX im Rahmen des DigiNet-Projektes sind als Beispiel zu verstehen und müssen projektspezifisch angepasst werden. Zunächst wird die Installation des E-PIX beschrieben. Hierbei liegt der Fokus auf einem von Windows betriebenen System. Danach erfolgt die Beschreibung der Konfiguration der E-PIX Domäne sowie die Umwandlung von identifizierenden Daten (IDATs) wie z.B. der Krankenversicherernummer zu Bloomfiltern.

Eine detaillierte Beschreibung zum E-PIX finden Sie im Handbuch unter:

<http://www.ths-greifswald.de/e-pix/handbuch>

Hintergründe, Definitionen, Konzept sowie Funktionalitäten des Dienstes sind umfangreich im Handbuch in den Kapiteln 1-5 beschrieben.

Die entsprechenden Kapitelangaben zum Handbuch innerhalb dieser Leitlinie dienen lediglich dazu, im Falle eines vertieften Interesses an der Software weiterführende Informationen bereitzustellen. Es ist nicht erforderlich, die im Folgenden angegebenen Kapitel zum Handbuch zu lesen. Für die Installation und Konfiguration des E-PIX im Rahmen des DigiNet-Projektes ist diese Leitlinie ausreichend.

Bei etwaigen Fragen zur Installation und Konfiguration des E-PIX können Sie sich gerne per Mail an die Unabhängige Treuhandstelle Greifswald ([kontakt-ths@uni-greifswald.de](mailto:kontakt-ths@uni-greifswald.de)) wenden oder das Kontaktformular der THS (<https://www.ths-greifswald.de/kontakt>) nutzen.

## 2. Installation des E-PIX unter „Docker für Windows“

Bitte beachten Sie, dass ein Nutzer mit Adminrechten zum Installieren und Starten der Software notwendig ist. Der E-PIX wird standardmäßig als Docker-Container betrieben. Für die Installation müssen Sie zunächst die aktuelle Version von Docker und Docker-Compose<sup>1</sup> auf dem Zielsystem installieren. Die Anleitungen dafür finden Sie unter folgenden Links:

<https://docs.docker.com/install>

<https://docs.docker.com/compose/install/>

Weitere Informationen zu den Systemanforderungen finden Sie im Kapitel 6.1 des Handbuchs.

Nachdem Sie die Docker Desktop App erfolgreich heruntergeladen und installiert haben, starten Sie diese auf Ihrem Zielsystem. Anschließend können Sie sich den E-PIX über die Webseite der THS unter folgenden Link herunterladen:

<https://www.ths-greifswald.de/forscher/e-pix/#download>

Entpacken Sie das Archiv an der gewünschten Stelle. Danach kann der E-PIX über Docker-Compose gestartet werden. Die notwendigen Schritte hierzu auf einem Windows-System sind im Folgendem beschrieben:

Starten Sie die Windows Console CMD mit Adminrechten und wechseln Sie in das gewählte Verzeichnis (enthält die Datei docker-compose.yml). Damit der E-PIX bei Windows problemlos gestartet werden kann, setzen Sie in der Datei envs/ttp\_commons.env den Parameter #WF\_MARKERFILES = AUTO auf FALSE und entfernen Sie die vorangehende Raute (#). Der Konfigurationsparameter sieht nun wie folgt aus:

```
„WF_MARKERFILES = FALSE“
```

Anhand folgenden Befehles in der Konsole können Sie nun den E-PIX über Docker-Compose starten:

```
„docker-compose up“
```

Das Starten der Software kann wenige Minuten in Anspruch nehmen. E-PIX wurde erfolgreich installiert, wenn Ihnen folgender Befehl angezeigt wird:

```
„Wildfly 26.1.2. Final [...] started in ...“
```

Sie können nun den E-PIX auf Ihrem Zielsystem verwenden. Öffnen Sie dafür einen neuen Tab in Ihrem Webbrowser und geben Sie folgende URL ein:

```
„localhost:8080/epix-web“
```

Weitere Informationen zum Download und Starten des Dienstes per Docker finden Sie im Kapitel 6.2<sup>2</sup>.

---

<sup>1</sup> Docker-Compose ist mittlerweile in der Docker Desktop App integriert, sodass nur eine App auf dem Computer installiert werden muss.

<sup>2</sup> Bitte beachten Sie, dass die Anweisungen zum Downloaden und Starten von E-PIX im Handbuch mit dem Computer-Betriebssystem Linux und **nicht** mit Windows beschrieben wurden.

### 3. Konfiguration des E-PIX via Oberfläche

#### E-PIX testen

Möchten Sie den E-PIX ausprobieren, können Sie hierfür die graphische Benutzeroberfläche auf dem Demo-System der THS verwenden. Sie finden diese unter folgenden Link:

<https://demo.ths-greifswald.de/epix-web/>

**Bitte geben Sie keine Echt-Daten ein**, da diese für alle Webseitenbesucher ersichtlich sind.

Das System wird täglich zurückgesetzt, sodass vorherige Eingaben gelöscht sind.

Der von Ihnen installierte E-PIX ist standardmäßig erreichbar unter <URL>“:8080/epix-web/“. Läuft der E-PIX auf Ihrem lokalen Gerät, kann dieser beispielsweise unter <http://localhost:8080/epix-web/> aufgerufen werden.

Der E-PIX verwaltet die eingegebenen Daten in Domänen. Daher muss zur Eingabe des IDAT wie z.B. der Krankenversichertennummern zunächst eine Domäne angelegt werden. Diese ermöglicht die Umwandlung des IDAT zu Bloomfiltern. Für jede Domäne muss eine sichere Quelle und eine Identifier-Domäne angegeben werden (vgl. Handbuch Kapitel 7.1). Der E-PIX hat eine Identifier-Domäne standardmäßig hinterlegt, welche für diesen Zweck verwendet werden kann (MPI) (vgl. Abb. 1).



| Name | Schlüssel | OID                        |
|------|-----------|----------------------------|
| MPI  | MPI       | 1.2.276.0.76.3.1.132.1.1.1 |

Abbildung 1 - Identifier-Domäne

#### 3.1 Anlegen der Domänen und Quellen

Klicken Sie im linken Menü unter „Einstellungen“ auf „Domänen/Quellen/Identifier“ (vgl. Abb. 2).

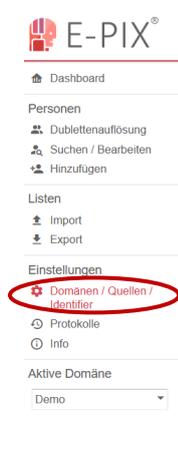
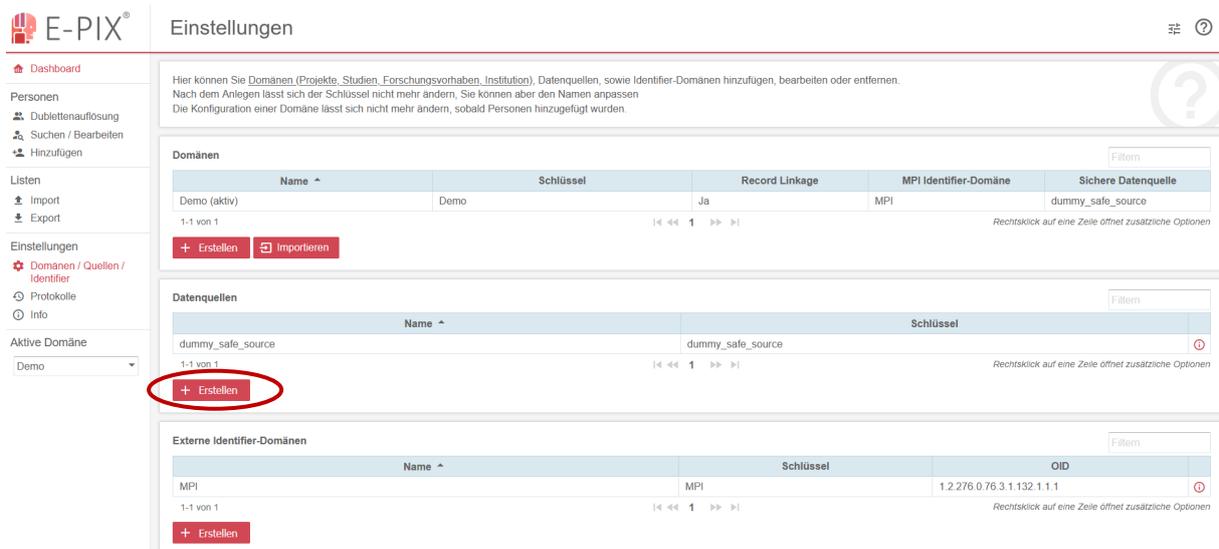


Abbildung 2 - Menü des E-PIX

## 3.2 Datenquelle anlegen (Sichere Quelle)

Klicken Sie anschließend unter „Datenquellen“ auf „Erstellen“ (vgl. Abb. 3).



The screenshot shows the 'Einstellungen' (Settings) page in E-PIX. The left sidebar contains navigation options like 'Dashboard', 'Personen', 'Listen', 'Einstellungen', and 'Aktive Domäne'. The main content area is titled 'Einstellungen' and contains three tables: 'Domänen', 'Datenquellen', and 'Externe Identifier-Domänen'. In the 'Datenquellen' table, the 'Erstellen' button is highlighted with a red circle.

| Name         | Schlüssel | Record Linkage | MPI Identifier-Domäne | Sichere Datenquelle |
|--------------|-----------|----------------|-----------------------|---------------------|
| Demo (aktiv) | Demo      | Ja             | MPI                   | dummy_safe_source   |

| Name              | Schlüssel         |
|-------------------|-------------------|
| dummy_safe_source | dummy_safe_source |

| Name | Schlüssel | OID                        |
|------|-----------|----------------------------|
| MPI  | MPI       | 1.2.276.0.76.3.1.132.1.1.1 |

Abbildung 3 - Oberfläche "Einstellungen" zum Erstellen einer Sicheren Quelle

Füllen Sie den Reiter „**Neue Datenquelle**“ nun wie folgt aus:

- **Name:** Beispiel „LKR“ bzw. „Krankenkasse“ (frei wählbar)
- **Schlüssel:** Beispiel „LKR“ bzw. „Krankenkasse“ (frei wählbar)
- **Beschreibung:** frei wählbar (Freitext; ggf. leer lassen)

(vgl. Handbuch Kapitel 7.1.1 (Erläuterung des Begriffs „Quelle“ in Kapitel 3 Begriffsbestimmungen))

Klicken Sie anschließend auf „**Erstellen**“ (vgl. Abb. 4).



The screenshot shows a modal window titled 'Neue Datenquelle' with a close button (X). It contains three input fields: 'Name \*' with the value 'LKR', 'Schlüssel' with the value 'LKR', and 'Beschreibung' which is empty. Below the description field, it says '255 Zeichen verbleibend'. At the bottom, there are two buttons: 'Erstellen' (with a checkmark) and 'Abbrechen' (with an X).

Abbildung 4 - Reiter "Neue Datenquelle" anlegen

Nach dem erfolgreichen Anlegen einer „**Sicheren Quelle**“ erfolgt das Anlegen der Domäne (vgl. Handbuch Kapitel 7.1.3).

### 3.3 Domäne anlegen (Matching-Domäne)

Für das Anlegen der Domäne gibt es zwei Möglichkeiten:

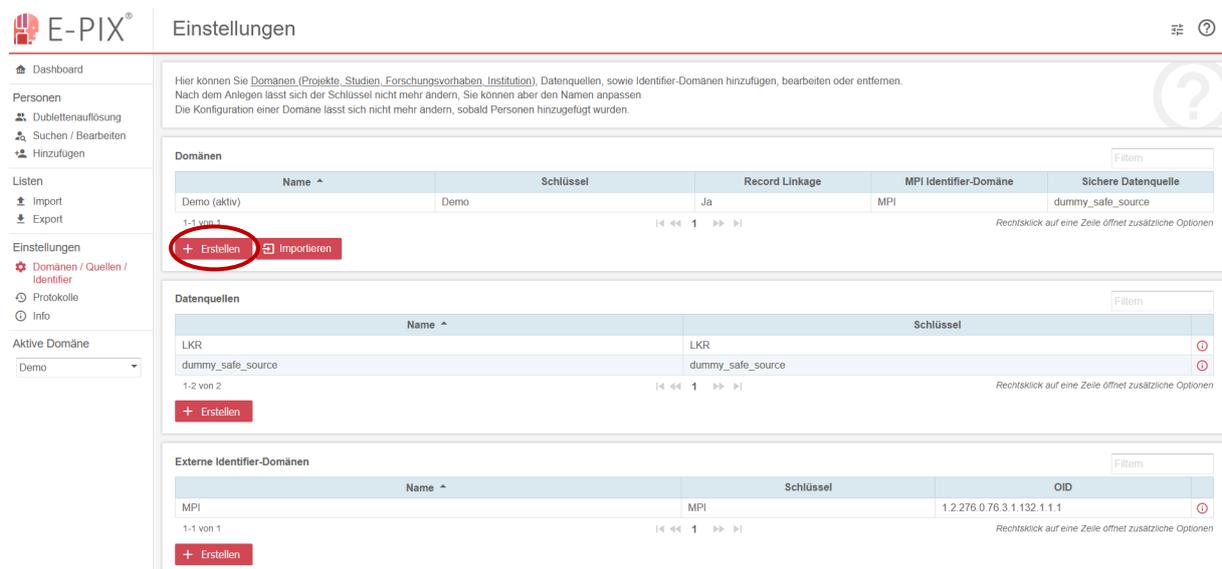
- 1) Die Domäne wird manuell gemäß der Konfiguration durch Klicken auf „Erstellen“ angelegt (siehe Kapitel 3.3.1 (Domäne manuell anlegen)).
- 2) Die Domäne wird mittels einer vorkonfigurierten XML-Datei durch Klicken auf „Importieren“ angelegt (siehe Kapitel 3.3.2 (Domäne anlegen via XML-Datei)).

Beide Vorgehensweisen werden nun im Folgenden ausführlich erläutert.

#### 3.3.1 Domäne manuell anlegen

Bitte beachten Sie, dass erst nach dem Ausfüllen aller Reiter der Domäne auf den Button „Hinzufügen“ zu klicken ist, um die Domäne anzulegen.

Klicken Sie unter „Domänen“ auf „Erstellen“ (vgl. Abb. 5).



The screenshot shows the 'Einstellungen' (Settings) page in the E-PIX system. The left sidebar contains navigation options like 'Dashboard', 'Personen', 'Listen', 'Einstellungen', and 'Aktive Domäne'. The main content area is titled 'Einstellungen' and contains three sections: 'Domänen', 'Datenquellen', and 'Externe Identifier-Domänen'. Each section has a table with columns for Name, Schlüssel, Record Linkage, MPI Identifier-Domäne, and Sichere Datenquelle. In the 'Domänen' section, the 'Erstellen' button is circled in red. Below the 'Domänen' table, there is a red 'Erstellen' button. The 'Datenquellen' and 'Externe Identifier-Domänen' sections also have 'Erstellen' buttons.

Abbildung 5 - Oberfläche "Einstellungen" zum Erstellen einer Domäne

#### 3.3.1.1 Anlegen einer Domäne: Reiter Einstellungen

Füllen Sie den Reiter „Einstellungen“ nun wie folgt aus (vgl. Abb. 6):

- **Name:** Beispiel „DigiNet\_nNGM\_Vergleichsgruppe“
- **Schlüssel:** Beispiel „DigiNet\_nNGM\_Vergleichsgruppe“  
(wird automatisch erzeugt, wenn das Feld frei gelassen wird)
- **Beschreibung:** Beispiel „Erzeugen eines Bloomfilters für die KV-Nr. der Patienten der bevölkerungsbezogenen Vergleichsgruppe“  
(Freitext; ggf. leer lassen)
- **Sichere Datenquelle:** „LKR“ bzw. „Krankenkasse“ (Zuvor angelegte Datenquelle)
- **Sende Benachrichtigung durch die Weboberfläche:** Keinen Haken setzen

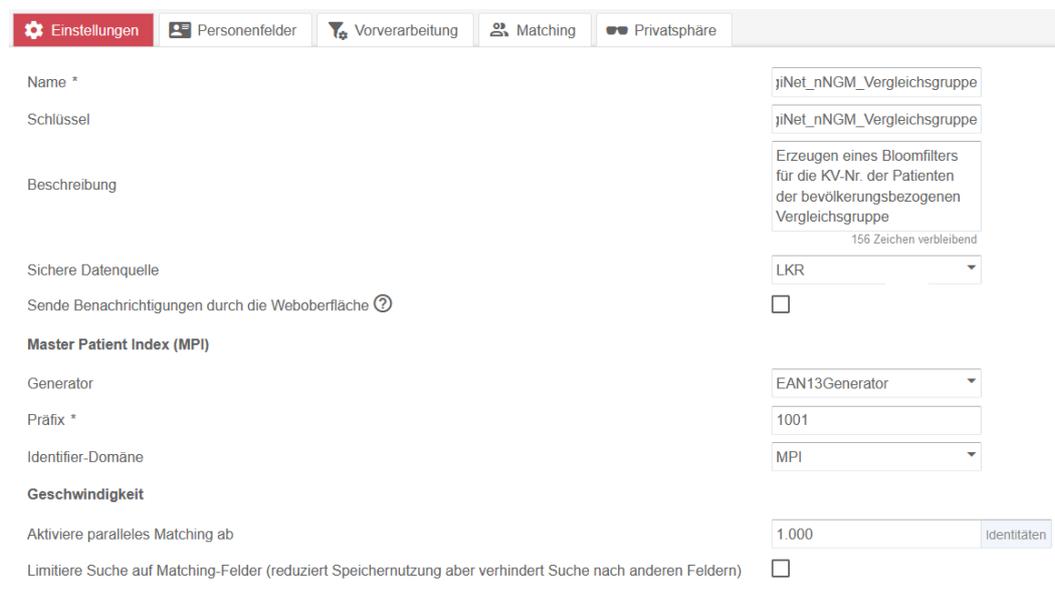
#### Master Patient Index (MPI)

- **Generator:** EAN13Generator
- **Präfix:** 1001
- **Identifizier-Domäne:** MPI

#### Geschwindigkeit

- **Aktiviere paralleles Matching ab:** 1.000
- **Limitiere Suche auf Matching-Felder (reduziert Speichernutzung aber verhindert Suche nach anderen Feldern):** Keinen Haken setzen

Weitere Informationen finden Sie im Handbuch Kapitel 7.1.3.1.



**Einstellungen** | Personenfelder | Vorverarbeitung | Matching | Privatsphäre

Name \*

Schlüssel

Beschreibung   
156 Zeichen verbleibend

Sichere Datenquelle

Sende Benachrichtigungen durch die Weboberfläche

**Master Patient Index (MPI)**

Generator

Präfix \*

Identifizier-Domäne

**Geschwindigkeit**

Aktiviere paralleles Matching ab  Identitäten

Limitiere Suche auf Matching-Felder (reduziert Speichernutzung aber verhindert Suche nach anderen Feldern)

Abbildung 6 - Reiter "Einstellungen" beim Anlegen einer Domäne

### 3.3.1.2 Anlegen einer Domäne: Reiter Personenfelder

Füllen Sie den Reiter „Personenfelder“ nun wie folgt aus:

- 1) Klicken Sie auf das Auswahlnü „Pflichtfelder“ und wählen Sie „Feld 1“ aus, indem Sie einen Haken setzen (vgl. Abb. 7).
- 2) Entfernen Sie alle weiteren Haken, die standardmäßig gesetzt sind.
- 3) Schließen Sie das Fenster wieder mit einem Klick auf die Oberfläche.

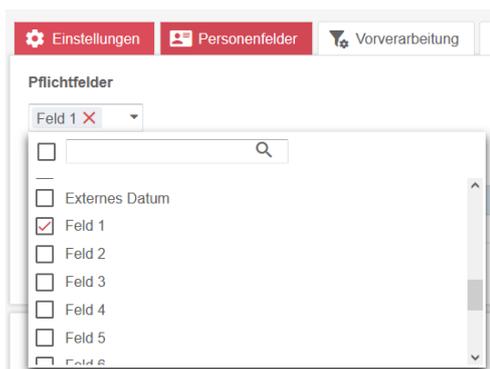


Abbildung 7 - Reiter "Personenfelder" zum Auswählen der Pflichtfelder

- 4) Klicken Sie nun unterhalb „Zusatzfelder“ auf „Zusatzfeld hinzufügen“.
- 5) Füllen Sie jeweils „Feld“ und „Bezeichnung“ aus (vgl. Abb. 8) und klicken Sie anschließend auf „Hinzufügen“.

Dies wiederholen Sie für alle Zusatzfelder. Tragen Sie dabei folgende Werte ein:

| Feld                  | Bezeichnung |
|-----------------------|-------------|
| Feld 1 (50 Zeichen)   | KVNR        |
| Feld 8 (1000 Zeichen) | Bloomfilter |

Zusatzfeld ×

Feld

Bezeichnung \*

Abbildung 8 - Reiter "Personenfelder" zum Anlegen der Zusatzfelder

Der Reiter „Personenfelder“ sollte nun wie folgt aussehen:



Abbildung 9 - Reiter "Personenfelder" zum Erstellen einer Domäne

Weitere Informationen finden Sie im Handbuch Kapitel 7.1.3.2.

### 3.3.1.3 Anlegen einer Domäne: Reiter Vorverarbeitung

Füllen Sie den Reiter „Vorverarbeitung“ nun wie folgt aus (vgl. Abb. 10):

Entfernen Sie alle Felder aus der Tabelle „Felder mit Vorverarbeitung“, indem Sie auf das Mülltonnensymbol klicken.

Der Reiter „**Vorverarbeitung**“ sollte nun wie folgt aussehen:



Abbildung 10 - Reiter "Vorverarbeitung" zum Erstellen einer Domäne

Weitere Informationen finden Sie im Handbuch Kapitel 7.1.3.3.

### 3.3.1.4 Anlegen einer Domäne: Reiter Matching

Füllen Sie den Reiter „**Matching**“ wie folgt aus (vgl. Abb. 12):

- **Führe Record Linkage durch:** Ja
- **Prüfe, ob übergebene identifizierende Daten neben der Person des Identifiers gleich gut mit anderen Personen:** Keinen Haken setzen
- **Mindestscore für mögliche Match:** 2.99
- **Mindestscore für automatische Match:** Haken bei „nie“ setzen

Entfernen Sie alle bereits vorher bestehenden „**Matching Felder**“ sowie „**Gründe für Dublettenauflösung**“, indem Sie auf das Mülltonnensymbol klicken.

Klicken Sie nun auf „**Matching Feld hinzufügen**“.

Füllen Sie nun das „**Matching Feld**“ wie folgt aus (vgl. Abb. 11):

- **Feld:** Feld 1
- **Algorithmus:** DeterministicAlgorithm
- **Modus:** Text
- **Blocking ab Übereinstimmung von:** 0%
- **Matching ab Übereinstimmung von:** 100%
- **Gewichtung:** 1.0

Klicken Sie anschließend auf „Speichern“.

**Matching Feld** ✕

Feld

Algorithmus

Modus

Blocking ab Übereinstimmung von \*

Matching ab Übereinstimmung von \*

Gewichtung \*

**Multi-Wert Feld**

Abbildung 11 - Reiter "Matching" zum Hinzufügen eines neuen Matching Feldes

Der Reiter „Matching“ sollte nun wie folgt aussehen:

**Matching**

Führe Record Linkage durch

Prüfe, ob übergebene identifizierende Daten neben der Person des Identifiers gleich gut mit anderen Personen matchen.

Mindestscore für möglichen Match \*

Mindestscore für automatischen Match \*    nie

**Matching Felder**

| Feld   | Algorithmus            | Modus | Blocking ab Übereinstimmung von | Matching ab Übereinstimmung von | Gewichtung |   |
|--------|------------------------|-------|---------------------------------|---------------------------------|------------|---|
| Feld 1 | DeterministicAlgorithm | Text  | 0 %                             | 100 %                           | 1.0        | <input type="button" value="✎"/> <input type="button" value="✖"/> |

[+ Matching Feld hinzufügen](#)

**Gründe für Dublettenauflösung**

| Bezeichnung                | Hinweis |
|----------------------------|---------|
| Keine Datensätze gefunden. |         |

[+ Grund hinzufügen](#)

Abbildung 12 - Reiter "Matching" zum Erstellen einer Domäne

Weitere Informationen finden Sie im Handbuch Kapitel 7.1.3.4.

### 3.3.1.5 Anlegen einer Domäne: Reiter Privatsphäre

Füllen Sie den Reiter „Privatsphäre“ nun wie folgt aus:

Klicken Sie unterhalb „Bloomfilter“ auf „Bloomfilter hinzufügen“.

Füllen Sie die Felder des „Bloomfilters“ wie folgt aus (vgl. Abb. 13)<sup>3</sup>:

- **Algorithmus:** RandomHashingStrategy
- **Alphabet:** ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789
- **Länge:** 300
- **Länge der N-Gramme:** 2
- **Bits pro N-Gramm:** 20
- **Anzahl der XOR-Faltungen:** 0
- **Balanced Bloomfilter:** Haken setzen
- **Seed:** x.xxx.xxx.xxx
- **Speicherfeld:** Feld 8

**Bloomfilter** ×

Algorithmus RandomHashingStrategy ▾

Alphabet \* ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

Länge \* 300 Bits

Länge der N-Gramme \* 2

Bits pro N-Gramm \* 20 Bits

Anzahl der XOR-Faltungen \* 0

Balanced Bloomfilter

Seed \* x.xxx.xxx.xxx

Speicherfeld \* Feld 8 ▾

**Quell-Felder**

| Feld   |  |
|--------|--|
| Feld 1 |  |

+ Quell-Feld hinzufügen

✓ Speichern
✗ Abbrechen

Abbildung 13 - Reiter "Privatsphäre" zum Hinzufügen eines Bloomfilters

<sup>3</sup> Der Seed entspricht einem individuell wählbaren numerischen Geheimcode. Ein identischer Seed ist Voraussetzung für die Reproduzierbarkeit identischer Bloomfilter.

Klicken Sie unterhalb „Quell-Felder“ auf „Quell-Feld hinzufügen“.

Füllen Sie das „Quell-Feld“ wie folgt aus (vgl. Abb. 14)<sup>4</sup>:

- **Feld:** Feld 1
- **Seed:** x.xxx.xxx.xxx

Klicken Sie anschließend auf „Speichern“.

**Quell-Feld** ×

Feld \*

Seed \*

Abbildung 14 - Reiter "Privatsphäre" zum Erstellen eines Quell-Feldes

Der „Bloomfilter“ sollte nun wie in Abbildung 13 dargestellt aussehen.

Klicken Sie nach dem Sie das Quell-Feld hinzugefügt und alle Felder korrekt ausgefüllt haben auf „Speichern“.

Der Reiter „Privatsphäre“ sollte nun wie in Abbildung 15 aussehen.

⚙️ Einstellungen
👤 Personenfelder
🔍 Vorverarbeitung
🔗 Matching
🔒 Privatsphäre

Speichermodus Identifizierende Daten und Bloomfilter ▾

**Bloomfilter**

| Algorithmus           | Quell-Felder | Speicherfeld |      |
|-----------------------|--------------|--------------|------|
| RandomHashingStrategy | Feld 1       | Feld 8       | ✎ 🗑️ |

+ Bloomfilter hinzufügen

---

Abbildung 15 - Reiter "Privatsphäre" zum Erstellen einer Domäne

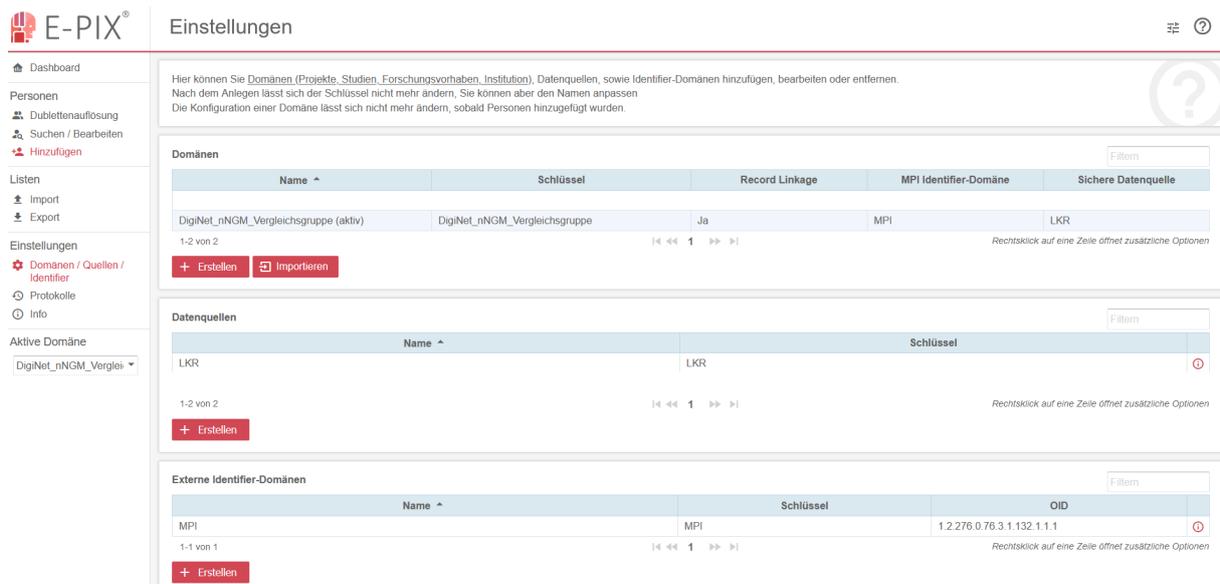
Weitere Informationen finden Sie im Handbuch Kapitel 7.1.3.5.

Die Einstellungen für die Domäne wurden nun festgelegt.

Klicken Sie abschließend auf „Hinzufügen“, um die Konfiguration der Domäne zu speichern (vgl. Abb. 15).

<sup>4</sup> Für das Quell-Feld wird zusätzlich ein Geheimcode bzw. Seed benötigt, der sich von dem Geheimcode bzw. Seed des Bloomfilters unterscheiden muss.

Die Oberfläche der Einstellungen des E-PIX sollte nun wie folgt aussehen:



The screenshot shows the 'Einstellungen' (Settings) page of the E-PIX system. The left sidebar contains navigation options: Dashboard, Personen (with sub-options: Dublettenauflösung, Suchen / Bearbeiten, Hinzufügen), Listen (with sub-options: Import, Export), Einstellungen (with sub-options: Domänen / Quellen / Identifier, Protokolle, Info), and Aktive Domäne (set to 'DigiNet\_nNGM\_Vergleis').

The main content area is titled 'Einstellungen' and includes a help icon. Below the title, there is a text box explaining that domains (Projects, Studies, Research Projects, Institutions), data sources, and identifier domains can be added, edited, or deleted. It notes that keys cannot be changed after creation, but names can be adjusted.

There are three main sections:

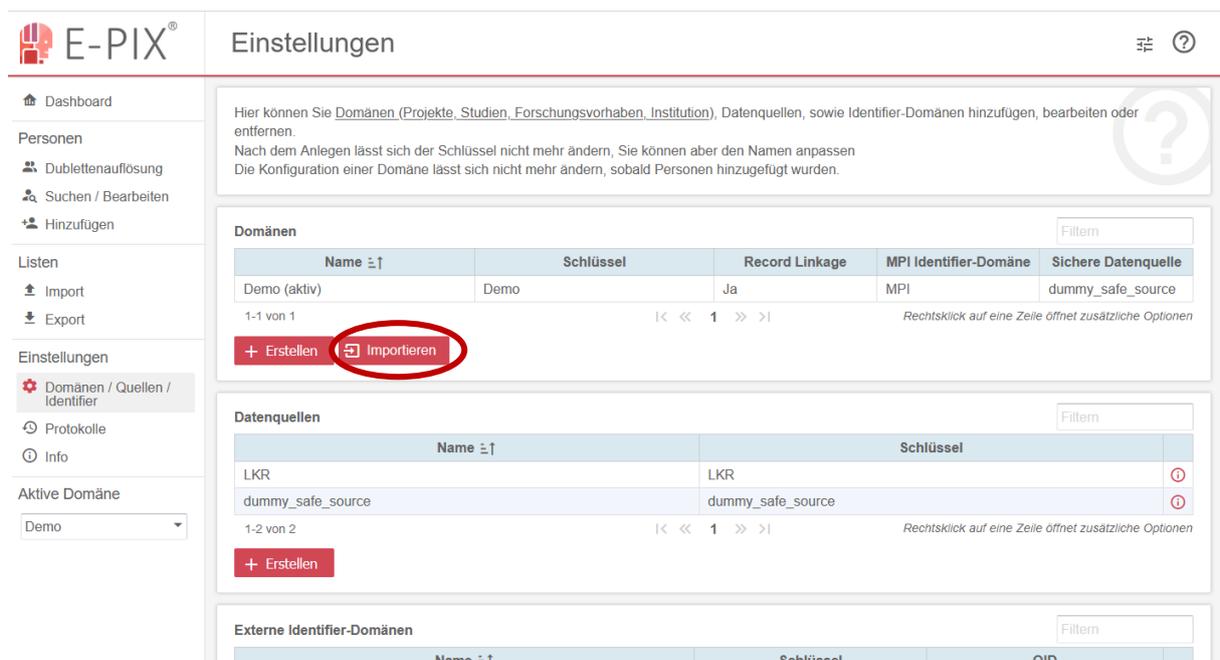
- Domänen (Domains):** A table with columns: Name, Schlüssel, Record Linkage, MPI Identifier-Domäne, and Sichere Datenquelle. One entry is visible: 'DigiNet\_nNGM\_Vergleichsgruppe (aktiv)' with key 'DigiNet\_nNGM\_Vergleichsgruppe', linkage 'Ja', MPI 'MPI', and source 'LKR'. Below the table are buttons for '+ Erstellen' and 'Importieren'.
- Datenquellen (Data Sources):** A table with columns: Name and Schlüssel. One entry is visible: 'LKR' with key 'LKR'. Below the table is a '+ Erstellen' button.
- Externe Identifier-Domänen (External Identifier Domains):** A table with columns: Name, Schlüssel, and OID. One entry is visible: 'MPI' with key 'MPI' and OID '1.2.276.0.76.3.1.132.1.1.1'. Below the table is a '+ Erstellen' button.

Abbildung 16 - Oberfläche "Einstellungen" des E-PIX nach erfolgter Konfiguration der Domäne

### 3.3.2 Domäne anlegen via XML-Datei

Seit Januar 2024 besteht die Möglichkeit der automatisierten Konfiguration mittels einer XML-Datei. Um diese Alternative zu nutzen, folgen Sie bitte allen Handlungsanweisungen bis einschließlich Kapitel 3.2. „Datenquelle anlegen (Sichere Quelle)“.

Nachdem Sie die „Sichere Datenquelle“ angelegt haben, können Sie die XML-Datei zur Konfiguration der Domäne importieren. Hierfür klicken Sie unter „Domänen“ auf „Importieren“. (vgl. Abb. 17).



This screenshot is similar to the previous one but highlights the 'Importieren' button in the 'Domänen' section with a red circle. The 'Aktive Domäne' dropdown is now set to 'Demo'. The 'Domänen' table shows a single entry: 'Demo (aktiv)' with key 'Demo', linkage 'Ja', MPI 'MPI', and source 'dummy\_safe\_source'. The 'Importieren' button is circled in red.

Abbildung 17 - Oberfläche "Einstellungen" zum Importieren einer Domäne

Suchen Sie in Ihrem Filesystem nach der XML-Datei die Sie von der Unabhängigen Treuhandstelle Greifswald per E-Mail erhalten haben.

Klicken Sie die XML-Datei an und anschließend auf „**Öffnen**“, um diese in die Domäne zu importieren.

Nach dem erfolgreichen Import der XML-Datei öffnet sich im nächsten Schritt die Konfiguration der Domäne automatisch mit den vorausgefüllten Reitern bzw. Feldern.

Folgende minimale Anpassungen müssen nun manuell vorgenommen werden:

Unter dem Reiter „**Einstellungen**“ muss der „**Name**“ der Domäne ergänzt werden (vgl. Abb. 18).

Prüfen Sie, dass die Angaben unter „**Sichere Datenquelle**“ sowie „**Identifizier-Domäne**“ korrekt ausgewählt wurden und passen Sie diese ggf. an (vgl. Abb. 18).

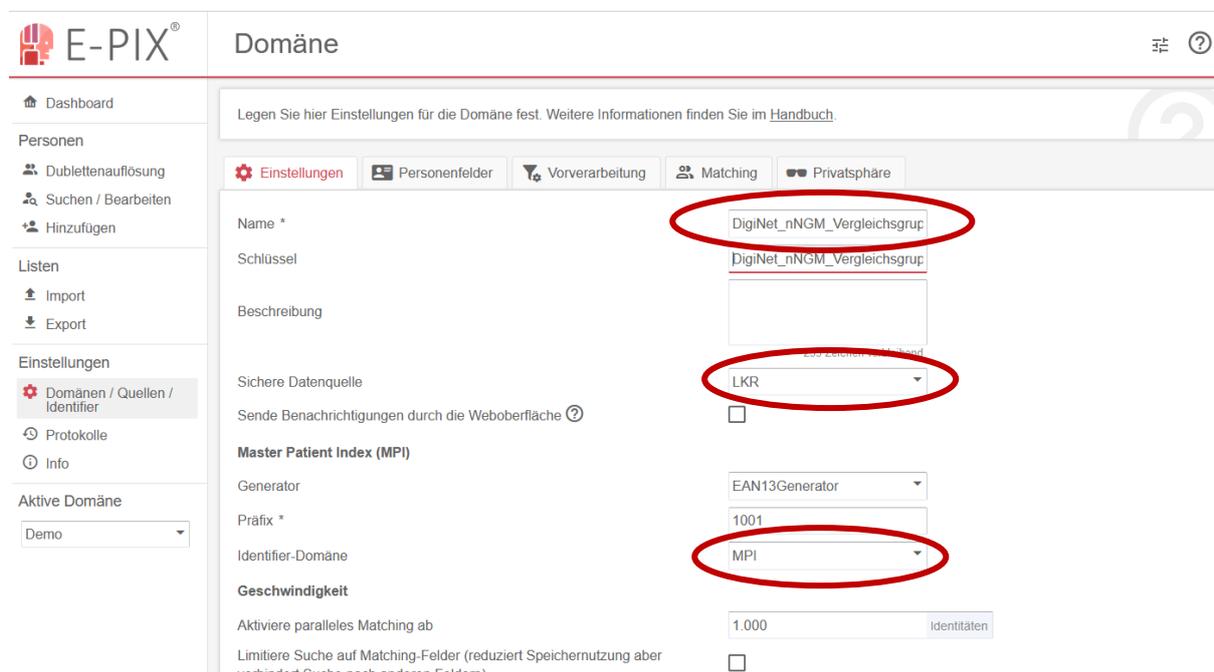
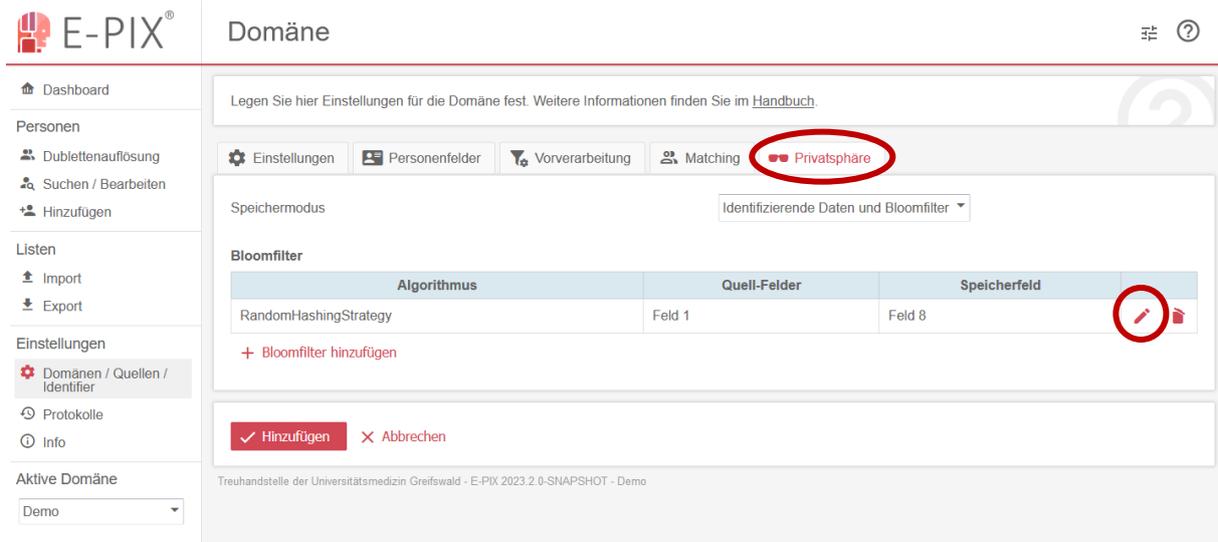


Abbildung 18 – Reiter „Einstellungen“ beim Anlegen einer Domäne

Die Reiter „**Personenfelder**“, „**Vorverarbeitung**“ und „**Matching**“ sind bereits vollständig ausgefüllt. Sie müssen keine Anpassungen innerhalb dieser drei Reiter vornehmen.

Um die beiden Geheimcodes (Seeds) für den Bloomfilter und das Quell-Feld zu ergänzen, gehen Sie bitte als Nächstes in den Reiter „**Privatsphäre**“. Klicken Sie dort auf das Symbol mit dem Stift, um den „**Bloomfilter**“ zu bearbeiten (vgl. Abb. 19).



E-PIX® Domäne

Legen Sie hier Einstellungen für die Domäne fest. Weitere Informationen finden Sie im [Handbuch](#).

Einstellungen | Personenfelder | Vorverarbeitung | Matching | **Privatsphäre**

Speichermodus: Identifizierende Daten und Bloomfilter

Bloomfilter

| Algorithmus           | Quell-Felder | Speicherfeld |
|-----------------------|--------------|--------------|
| RandomHashingStrategy | Feld 1       | Feld 8       |

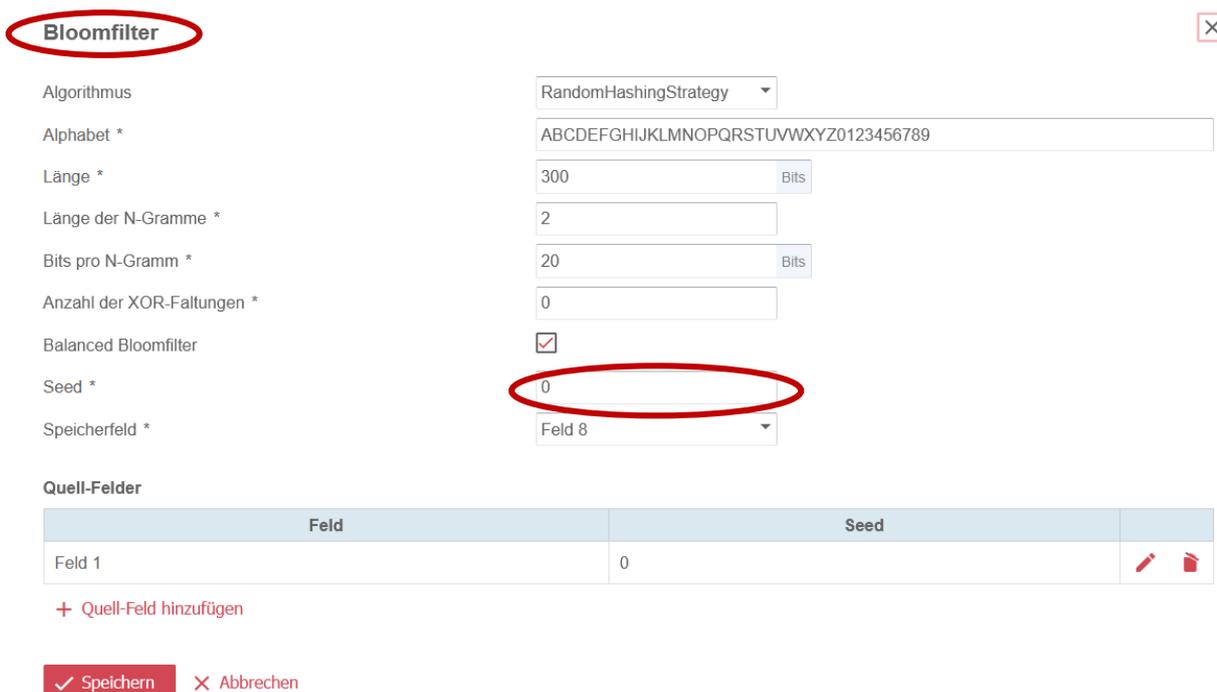
+ Bloomfilter hinzufügen

Hinzufügen | Abbrechen

Treuhandstelle der Universitätsmedizin Greifswald - E-PIX 2023.2.0-SNAPSHOT - Demo

Abbildung 19 - Reiter "Privatsphäre" beim Anlegen einer Domäne

Es öffnet sich ein Fenster, in dem Sie nun den Seed für den „Bloomfilter“ eingeben können (vgl. Abb. 20)<sup>5</sup>.



**Bloomfilter**

Algorithmus: RandomHashingStrategy

Alphabet \*: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

Länge \*: 300 Bits

Länge der N-Gramme \*: 2

Bits pro N-Gramm \*: 20 Bits

Anzahl der XOR-Faltungen \*: 0

Balanced Bloomfilter:

Seed \*: 0

Speicherfeld \*: Feld 8

Quell-Felder

| Feld   | Seed |
|--------|------|
| Feld 1 | 0    |

+ Quell-Feld hinzufügen

Speichern | Abbrechen

Abbildung 20 - Reiter "Privatsphäre" zum Hinzufügen eines Bloomfilters

**Hinweis:** Durch das Importieren der XML-Datei wird der Seed des Bloomfilters automatisch auf „0“ gesetzt, sodass Sie an dieser Stelle den entsprechenden Wert ergänzen müssen.

<sup>5</sup> Der Seed entspricht einem individuell wählbaren numerischen Geheimcode. Ein identischer Seed ist Voraussetzung für die Reproduzierbarkeit identischer Bloomfilter.

Um den Seed für das „Quell-Feld“ zu ergänzen, klicken Sie bitte unter „Quell-Felder“ auf das Symbol mit dem Stift (vgl. Abb. 21).

**Bloomfilter** ✕

Algorithmus: RandomHashingStrategy

Alphabet \*: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

Länge \*: 300 Bits

Länge der N-Gramme \*: 2

Bits pro N-Gramm \*: 20 Bits

Anzahl der XOR-Faltungen \*: 0

Balanced Bloomfilter:

Seed \*: 0

Speicherfeld \*: Feld 8

**Quell-Felder**

| Feld   | Seed |   |
|--------|------|---|
| Feld 1 | 0    |   |

+ Quell-Feld hinzufügen

Abbildung 21 - Reiter "Privatsphäre" zum Erstellen eines Quell-Feldes

Es öffnet sich ein Fenster, in dem Sie nun den Seed für das „Quell-Feld“ eingeben können und anschließend auf „Speichern“ klicken (vgl. Abb. 22)<sup>6</sup>.

**Bloomfilter** ✕

Algorithmus: RandomHashingStrategy

Alphabet \*: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

Länge \*: 300 Bits

Länge der N-Gramme \*: 2

Bits pro N-Gramm \*: 20 Bits

Anzahl der XOR-Faltungen \*: 0

Balanced Bloomfilter:

Seed \*: 0

Speicherfeld \*: Feld 8

**Quell-Felder**

| Feld   | Seed |   |
|--------|------|---|
| Feld 1 | 0    |   |

+ Quell-Feld hinzufügen

Abbildung 22 - Reiter "Privatsphäre" zum Erstellen eines Quell-Feldes

**Hinweis:** Durch das Importieren der XML-Datei wird der Seed des Quell-Feldes automatisch auf „0“ gesetzt, sodass Sie an dieser Stelle den entsprechenden Wert ergänzen müssen.

<sup>6</sup> Für das Quell-Feld wird zusätzlich ein Geheimcode bzw. Seed benötigt, der sich von dem Geheimcode bzw. Seed des Bloomfilters unterscheiden muss.

Nachdem Sie die beiden Geheimcodes (Seeds) ergänzt haben, klicken Sie bitte anschließend im Fenster „Bloomfilter“ auf „Speichern“ (vgl. Abb. 23).

### Bloomfilter ✕

Algorithmus: RandomHashingStrategy

Alphabet \*: ABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789

Länge \*: 300 Bits

Länge der N-Gramme \*: 2

Bits pro N-Gramm \*: 20 Bits

Anzahl der XOR-Faltungen \*: 0

Balanced Bloomfilter:

Seed \*: X.XXX.XXX.XXX

Speicherfeld \*: Feld 8

#### Quell-Felder

| Feld   | Seed  |     |
|--------|---|-----|
| Feld 1 | <span style="border: 2px solid red; border-radius: 50%; padding: 2px;">X.XXX.XXX.XXX</span> | ✎ 🗑 |

+ Quell-Feld hinzufügen

✓ Speichern
✕ Abbrechen

Abbildung 23 - Reiter "Privatsphäre" zum Hinzufügen eines Bloomfilters

Klicken Sie abschließend auf „Hinzufügen“ (vgl. Abb. 24). Die Domäne ist nun erfolgreich erstellt.



## Domäne ☰ ?

Legen Sie hier Einstellungen für die Domäne fest. Weitere Informationen finden Sie im [Handbuch](#).

⚙️ Einstellungen
👤 Personenfelder
🔍 Vorverarbeitung
👥 Matching
❤️ Privatsphäre

Speichermodus: Identifizierende Daten und Bloomfilter

| Algorithmus           | Quell-Felder | Speicherfeld |     |
|-----------------------|--------------|--------------|-----|
| RandomHashingStrategy | Feld 1       | Feld 8       | ✎ 🗑 |

+ Bloomfilter hinzufügen

✓ Hinzufügen
✕ Abbrechen

Aktive Domäne: Demo

Treuhandstelle der Universitätsmedizin Greifswald - E-PIX 2023 2.0-SNAPSHOT - Demo

Abbildung 24 - Reiter "Privatsphäre" zum Anlegen einer Domäne

## 4. Umwandlung der IDAT zu Bloomfiltern

**Hinweis:** Bitte beachten Sie, dass keine IDATs importiert werden sollten, solange die Konfiguration nicht abgeschlossen ist. Andernfalls müssten Sie den E-PIX neu starten und konfigurieren.

Nach dem Erstellen der Sicheren Quelle und der (Matching-)Domäne können Sie nun das entsprechende IDAT wie z.B. die Krankenversicherungsnummer zu Bloomfiltern umwandeln. Hierzu wird eine CSV-Datei mit dem IDAT in den E-PIX importiert. Im Weiteren wird davon ausgegangen, dass die CSV-Datei nur eine Spalte mit dem entsprechenden IDAT enthält. Der E-PIX generiert im Anschluss eine CSV-Datei mit dem IDAT und den daraus erzeugten Bloomfiltern.

Für das Importieren der CSV-Dateien gehen Sie bitte wie folgt vor:

- 1) Klicken Sie im linken Menü unter „Listen“ auf „Import“ (vgl. Abb. 25).

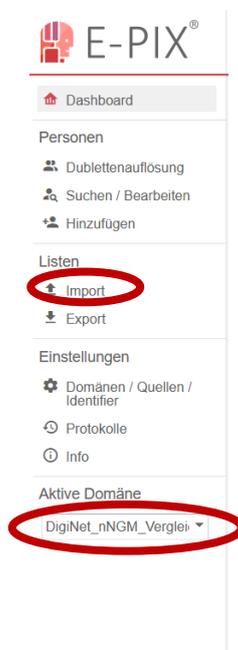


Abbildung 25 - Menü des E-PIX

**Hinweis:** Wenn Sie im Demo-System arbeiten oder bereits mehrere Domänen angelegt haben, dann achten Sie bitte darauf, dass unter „Aktive Domäne“ die von Ihnen angelegte Domäne ausgewählt ist (vgl. Abb. 25).

- 2) Setzen Sie den Haken bei „Die Liste besitzt eine Kopfzeile mit Spaltenbezeichnungen“, wenn die von Ihnen hochgeladene CSV-Datei eine Kopfzeile bzw. Spaltenbezeichnung besitzt (vgl. Abb. 26).

- 3) Klicken Sie auf „Datei auswählen“ und laden Sie die CSV-Datei hoch.

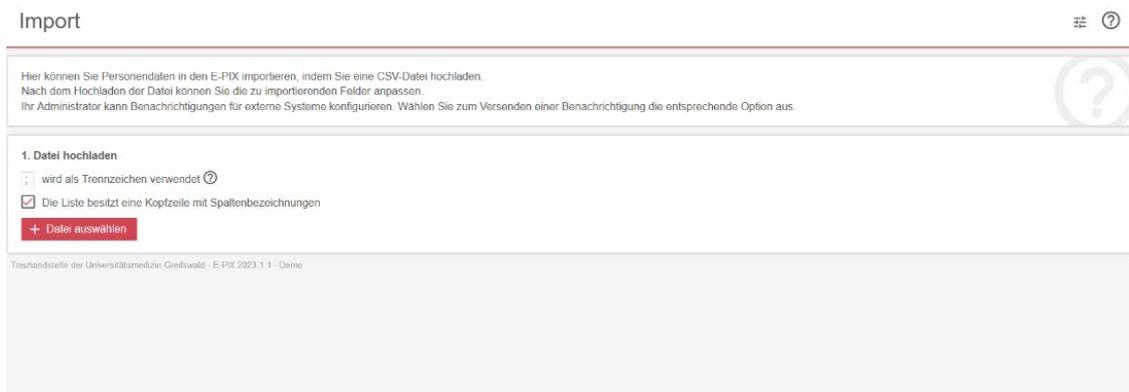


Abbildung 26 - Import über die Oberfläche

Unter „Felder auswählen und Typ festlegen“ sehen Sie nun den ersten Datensatz aus der Datei (vgl. Abb. 27)<sup>7</sup>.

- 4) Setzen Sie einen Haken bei allen Feldern, die importiert werden sollen.  
5) Wählen Sie eine Datenquelle aus. Dies entspricht Ihrer „Sicheren Quelle“.  
6) Klicken Sie abschließend auf „Importieren“.

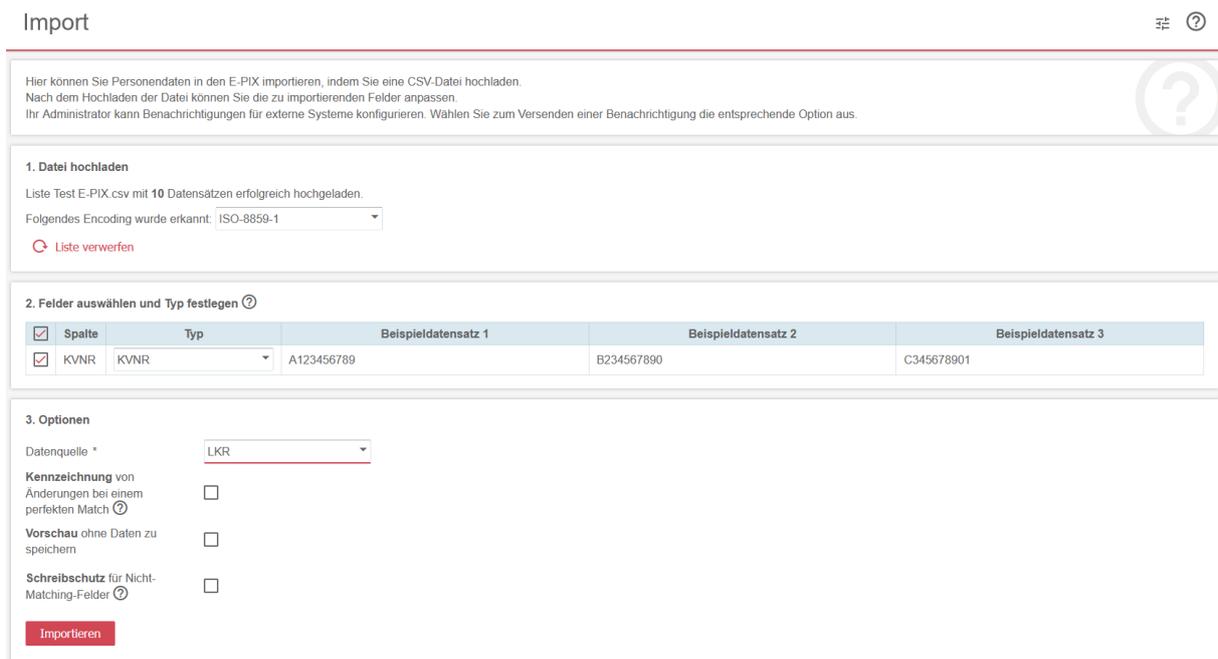


Abbildung 27 - Importieren von Daten

Nach erfolgreichem Import (vgl. Kapitel 4) können Sie das Ergebnis herunterladen, indem Sie auf „CSV herunterladen“ klicken. Sie erhalten eine CSV-Datei mit dem IDAT und den erzeugten Bloomfiltern sowie weiteren Variablen (Matchstatus, MPI, Hinzugefügt am, Letzte Bearbeitung um) (vgl. Abb. 28)<sup>7</sup>.

Weitere Informationen finden Sie im Handbuch Kapitel 7.8.

<sup>7</sup> Die hier dargestellten Krankenversicherungsnummern sind frei erfunden und können keinen realen Personen zugeordnet werden.

Erfolgreiche Importe: 10

| Matchstatus | MPJ           | KVNR        | Bloomfilter                                  | Hinzugefügt am      | Letzte Bearbeitung am |
|-------------|---------------|-------------|--|---------------------|-----------------------|
| NO_MATCH    | 1001000000059 | A123456789  | lucYybSfhN1gfl3xpYIIDU7dvQPovkdx4oW [...]    | 08.09.2023 13:54:00 | 08.09.2023 13:54:00   |
| NO_MATCH    | 1001000000066 | B234567890  | kk+QyeQfhN0kXn1hfh+IQBU7IuwPpvkclotG [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000073 | C345678901  | i6+QzbQxd0IX31xfheiQBS6cP9f4N14AMIX [...]    | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000080 | D456789012  | lqsczbQ9IN1kXj17ehPIQhU6crdf4NI5jMoU [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000097 | E567890123  | lrtezbQ9IN9kXn16WAOIQFUadrU3oNI5joIW [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000103 | F678901234  | kr9YzbQ/hN1mXn1xfA2oQZVaVvUbgNI5hooW [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000110 | G789012345  | Bq0YzbQ7rP1kXi1wfhWoRRH6VvUfgNI7po4W [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000127 | H890123456  | hq0YjbQm7NVkzrn1xfheiVBE4ZvU/oNhh7ho4e [...] | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000134 | I9012345671 | lq8azLQ3hN1wDn1hfhASBWyZnWfoJh5hooe [...]    | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |
| NO_MATCH    | 1001000000141 | J012345678  | lucYTbS/pN10Hn3x+pYASDUydvUfqJl5xqoW [...]   | 08.09.2023 13:54:01 | 08.09.2023 13:54:01   |

CSV herunterladen

Abbildung 28 - Oberfläche "Import" zum Herunterladen der CSV-Datei

## 5. Exportieren der IDAT und Bloomfilter via Export-Funktion

Wenn Sie die IDATs wie z.B. die Krankenversichertennummer (KVNr) wie in Kapitel 4 beschrieben bereits importiert haben und nun zeitversetzt die Zuordnung IDAT – Bloomfilter erhalten möchten, nutzen Sie dazu die Export-Funktion des E-PIX.

Dafür klicken Sie im Menü des E-PIX unter „Listen“ auf „Export“ (vgl. Abb. 29)

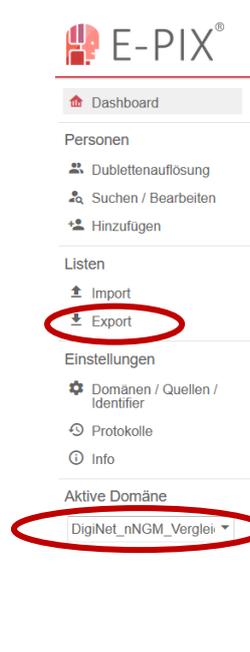
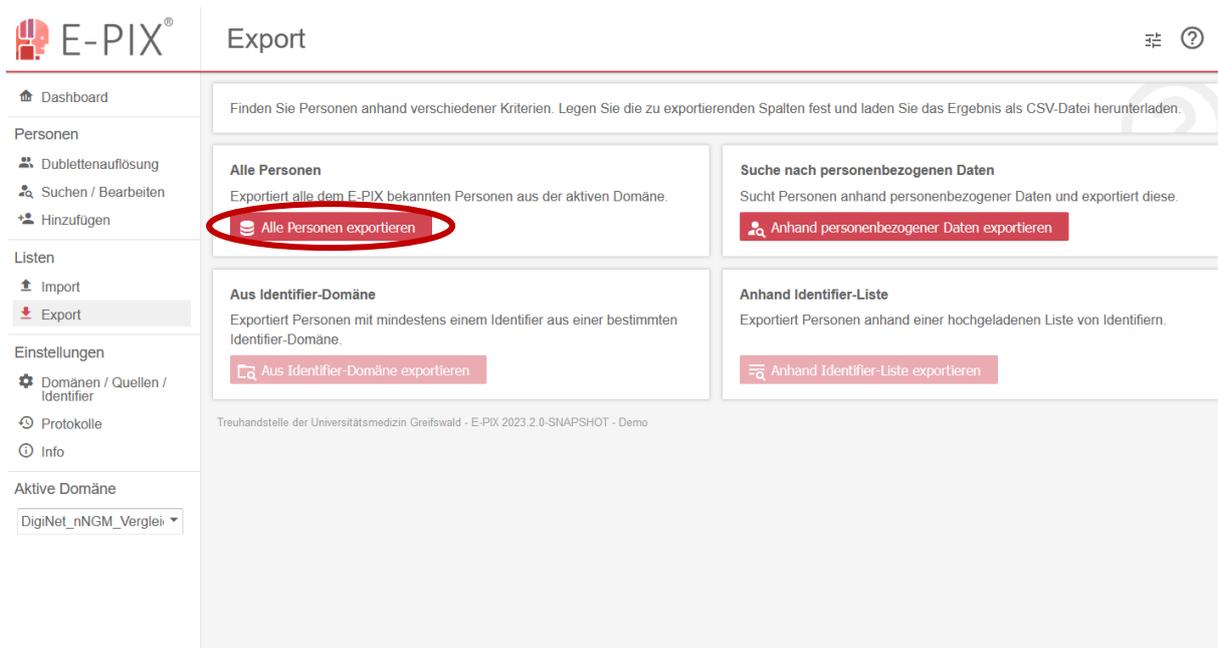


Abbildung 29 - Menü des E-PIX

**Hinweis:** Wenn Sie im Demo-System arbeiten oder bereits mehrere Domänen angelegt haben, dann achten Sie bitte darauf, dass unter „Aktive Domäne“ die von Ihnen angelegte Domäne ausgewählt ist (vgl. Abb. 29).

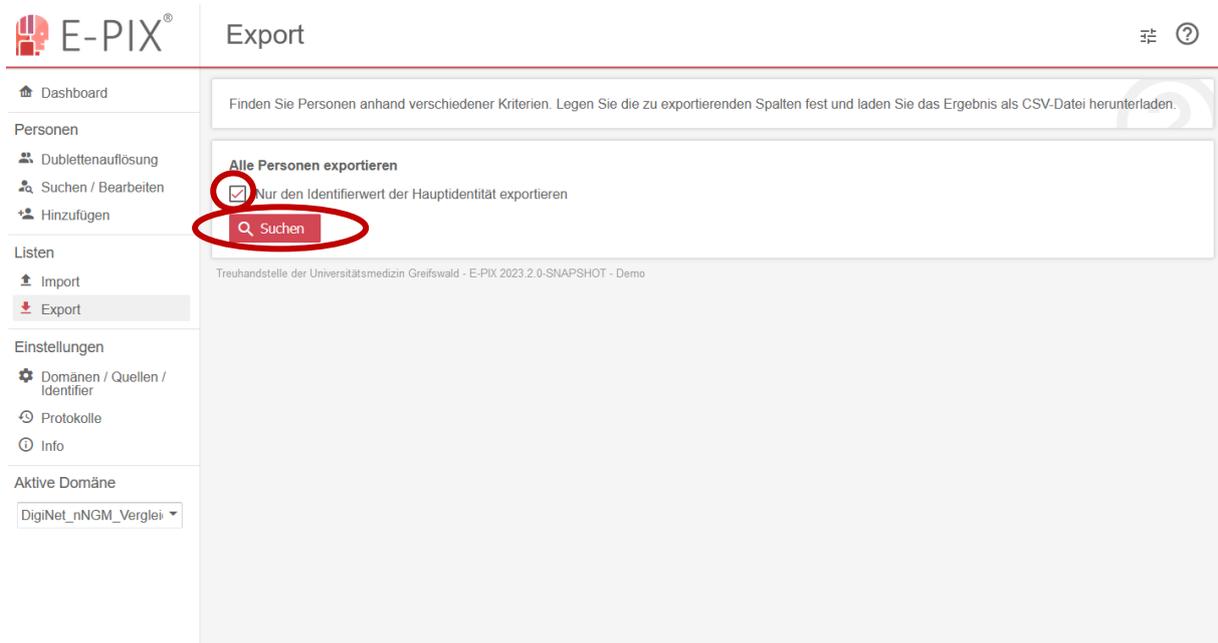
Klicken Sie jetzt auf „Alle Personen exportieren“ (vgl. Abb. 30).



The screenshot shows the E-PIX Export interface. On the left is a navigation menu with categories: Dashboard, Personen (Dublettenauflösung, Suchen / Bearbeiten, Hinzufügen), Listen (Import, Export), Einstellungen (Domänen / Quellen / Identifier, Protokolle, Info), and Aktive Domäne (DigiNet\_nNGM\_Verglei). The main content area is titled 'Export' and contains a search bar and four export options: 'Alle Personen exportieren' (circled in red), 'Suche nach personenbezogenen Daten', 'Aus Identifier-Domäne', and 'Anhand Identifier-Liste'. A footer note reads: 'Treuhandstelle der Universitätsmedizin Greifswald - E-PIX 2023 2.0-SNAPSHOT - Demo'.

Abbildung 30 - Exportieren von Daten

Belassen Sie den Hacken für das Item „Nur den Identifierwert der Hauptidentität exportieren“ und klicken Sie auf „Suchen“ (vgl. Abb. 31).



This screenshot shows the same E-PIX Export interface as in Abbildung 30, but with the 'Alle Personen exportieren' section expanded. It shows a checkbox labeled 'Nur den Identifierwert der Hauptidentität exportieren' which is checked, and a red 'Suchen' button below it, both of which are circled in red. The rest of the interface, including the navigation menu and footer, remains the same.

Abbildung 31 - Exportieren von Daten: Schritt "Alle Personen exportieren"

Klicken Sie nun auf „Leere Spalten ausschließen“ und anschließend auf das Kreuz in der Spalte „MPI“, um auch diese Variable aus dem Ergebnis zu entfernen (vgl. Abb. 32).

Abbildung 32 - Exportieren von Daten: Schritt „Exportspalten anpassen“

**Hinweis:** Die ausgegrauten Spalten werden nicht exportiert. Falls Sie nur das Feld 8 (Bloomfilter) als Variable exportieren möchten, können Sie das Feld 1 (KVNr) zusätzlich als Variable aus dem Ergebnis entfernen, indem Sie auch hier auf das Kreuz in der Spalte „Feld 1“ klicken.

Mittels der horizontalen Scroll-Leiste können Sie die beiden verbliebenen Variablen Feld 1 (KVNr) und Feld 8 (Bloomfilter) sehen, welche als CSV-Datei exportiert werden (vgl. Abb. 33).

Klicken Sie auf „CSV herunterladen“ und die Datei wird in Ihrem Download-Ordner gespeichert (vgl. Abb. 33).

Abbildung 33 - Exportieren von Daten: Schritt „CSV herunterladen“

Weitere Informationen finden Sie im Handbuch Kapitel 7.7.

Die THS Werkzeuge wie der E-PIX finden in der Praxis zunehmend Anwendung. Um den gemeinsamen Austausch zwischen den Projektpartnern bzw. Anwendern der Werkzeuge zu fördern, hat die Unabhängige Treuhandstelle der Universitätsmedizin Greifswald die THS Community als zentrale Anlaufstelle für spezifische Fragen initiiert. Weitere Informationen zur THS Community erhalten Sie unter folgenden Link:

<https://www.ths-greifswald.de/community/>

Hier können Sie sich bei Interesse auch zum nächsten THS Community Dialog direkt anmelden.