

sormas

# Datenbereinigung und -abgleich bei Import aus SurvNet zu SORMAS

Version 1.1.0

Stand 28.05.2021

sormas

## Wahrung der Vertraulichkeit

Dieses Dokument darf ohne schriftliche Genehmigung des Helmholtz Zentrum für Infektionsforschung weder ganz noch teilweise dupliziert, an Dritte weitergegeben oder anderweitig veröffentlicht werden. Dies gilt nicht für Kopien, die für die interne Verwendung bestimmt sind.

## Versionshistorie

Datum/ Uhrzeit	Beschreibung	Kürzel Autor	Version
06.05.2021	Dokument erstellt für SORMAS >= 1.59	TPA	1.0.0
07.05.2021	Behebung Syntaxfehler	DMA	1.0.1
28.05.2021	Anpassung Ziel des automatischen Abgleichs	DMA	1.1.0

## Inhaltsverzeichnis

Wahrung der Vertraulichkeit.....	2
Versionshistorie.....	2
Einleitung.....	3
Inbetriebnahme der Schnittstellen .....	3
Ablauf der Inbetriebnahme der Schnittstellen aus technischer Sicht.....	3
Automatischer Datenabgleich im Rahmen des Imports .....	4
Erkennung von Doubletten .....	4
Regeln der Zusammenführung.....	4
Manuelle Datenbereinigung vor der Produktivsetzung .....	4
Manueller Datenabgleich nach erfolgtem Import .....	4
Unterstützung bei dem Datenabgleich durch das Projekt .....	5
Voraussetzung .....	5
Bereitstellung des Teams .....	5

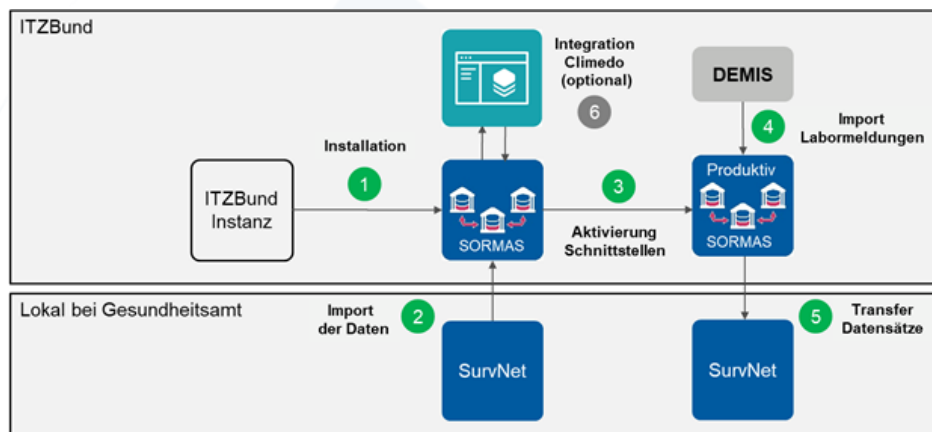
## Einleitung

### Inbetriebnahme der Schnittstellen

Die Inbetriebnahme der Schnittstellen erfolgt in enger Abstimmung mit dem Gesundheitsamt, da zeitnah der initiale Import und die Umstellung der Abläufe im Gesundheitsamt erfolgen. Im Folgenden ist der Ablauf der Migration aus technischer Sicht daher noch einmal dargestellt.

### Ablauf der Inbetriebnahme der Schnittstellen aus technischer Sicht

Das folgende Schaubild stellt den Ablauf der technischen Migration als Übersicht dar.



**Abbildung 4:** Ablauf der technischen Migration (schematisch)

Erläuterung zu Abbildung 4:

- Installation und Migration der Daten
  1. Einrichtung einer SORMAS-Instanz bei ITZBund durch Netzlink
  2. Migration der SORMAS-Bestandsdaten (falls vorhanden) im Auftrag des Gesundheitsamtes
  3. Übergabe der SORMAS-Instanz an das Gesundheitsamt
- Import der Daten
  1. Datenexport aus SurvNet und Import in SORMAS
  2. Automatischer Datenabgleich im Rahmen des Imports
  3. Manuelle Nachbereitung zur Datenbereinigung
- Aktivierung Schnittstellen
  1. Einrichtung SurvNet-Schnittstelle
  2. Einrichtung DEMIS-Schnittstelle
  3. Die tatsächliche Freischaltung der Schnittstellen erfolgt erst nach vollständigem Datenabgleich
- Transfer der Datensätze
  1. Übertragung von Fällen, Kontakten und Ereignissen an SurvNet
  2. Meldung an die Landesstelle aus SurvNet heraus
- Optional Inbetriebnahme von Climedo
  1. Einrichtung von Climedo und
  2. Freischaltung der Schnittstelle im Auftrag des Gesundheitsamtes

## Automatischer Datenabgleich im Rahmen des Imports

Das Ziel des automatischen Datenabgleichs ist es, doppelte Datensätze zu identifizieren und automatisch zusammenzuführen. Dafür werden beim Import aus SurvNet zunächst die Aktenzeichenfelder und die ExternalID aus SORMAS gelöscht.

### Erkennung von Doubletten

Für einen gegebenen SurvNet Fall „Surv1“ wird beim Import durch den SurvNet Connector in SORMAS folgende Logik zur Duplikatserkennung benutzt:

- ein Fall in SORMAS hat den gleichen Vornamen, Nachnamen, Geburtstag, Geschlecht, Krankheit, Meldedatum (+/- X Tage) und Landkreis (district). Felder, die leer sind, matchen dabei immer.

Im Importer wird „Surv1“ nur mit einem gefundenen Duplikat zusammengeführt, wenn eine eindeutige Zuordnung gegeben ist.

D.h. wenn eine Doublette nicht eindeutig bestimmt werden kann, muss in SORMAS die manuelle Duplikatserkennung und das Zusammenführen, bzw. Nicht-Zusammenführen erfolgen.

### Regeln der Zusammenführung

Wenn eine Doublette erkannt wird, dann wird sie automatisch entsprechend vorgegebener Regel zusammengeführt. Dabei wird in der Regel so vorgegangen, dass

1. Das Meldedatum immer aus der meldenden IfSG-Anwendung eingesetzt wird
2. Das Aktenzeichen von SurvNet immer das Feld Aktenzeichen von SORMAS überschreiben wird
3. Das Feld der Externen ID in SORMAS mit der Anwendungsübergreifenden GUID gefüllt wird
4. Felder, die in SurvNet leer sind, jedoch in SORMAS gefüllt sind, nicht verändert werden
5. Bei allen anderen Feldern, bei denen eine Entscheidung getroffen werden muss, der Inhalt von SurvNet als führend angesehen und der Wert der SORMAS Einträge verworfen wird

## Manuelle Datenbereinigung vor der Produktivsetzung

Die Daten von SurvNet und SORMAS könnten beispielsweise über einen Export zu Excel (bitte beachten Sie Hinweise zu UTF-8 auf der OEGD Seite) o.Ä. überführt und dort auf Gleichheit überprüft werden. So können Abweichungen identifiziert und gelöst werden.

## Manueller Datenabgleich nach erfolgtem Import

Das Ziel des manuellen Datenabgleichs ist es, doppelte Datensätze in SORMAS zu bereinigen und entweder zusammenzuführen oder bewusst nicht zusammenzuführen.

SORMAS ist mit einer Funktion ausgestattet, die es erlaubt, Doubletten automatisch zu erkennen und zusammenzuführen. Diese wird bei der manuellen Bereinigung genutzt.

### Unterstützung bei dem Datenabgleich durch das Projekt

Netzlink und HZI können Sie auf Anfrage bei dem manuellen Datenabgleich personell unterstützen. Dies bezieht sich aktuell ausschließlich auf die Datenbereinigung in SORMAS nach dem initialen Import.

#### Voraussetzung

Eine Voraussetzung, damit dies überhaupt möglich ist, ist die Unterzeichnung eines entsprechenden Vertrags, der dem Projektteam erlaubt, im Auftrag des Gesundheitsamtes die Daten in SORMAS zu sehen und zu bereinigen.

Dieser Vertrag wird mit Netzlink, dem HZI und Gesundheitsamt jeweils individuell abgeschlossen.

#### Bereitstellung des Teams

Zum erwarteten Abschluss des initialen Imports steht das Team zur Bereinigung der Daten bereit.

Der Zugriff auf die jeweilige SORMAS-Instanz erfolgt aus der gesicherten Umgebung von Netzlink heraus.

Für die Zeit der Bereinigung ist das IP-Whitelisting so eingestellt, dass der Zugriff auf die SORMAS-Instanz möglich ist.

1.