Fachkonzept

XFIM

Versionsübergreifende Übernahme von Daten aus FIM-Datenschemata

Scopeland Technology GmbH
Thomas Wolf, Geschäftsführer
Stand Oktober 2026, Entwurf

Datenschemata im FIM-Baukasten werden fortlaufend weiterentwickelt und liegen deshalb in mehreren Versionen vor. Behörden arbeiten nicht zwingend mit derselben Version. Eine Austauschdatei, die gegen eine andere Version erstellt wurde als die eigene, wird bei strenger Prüfung heute vollständig abgelehnt, obwohl der weit überwiegende Teil ihrer Inhalte unverändert verwendbar wäre. Dieses Fachkonzept beschreibt, wie XFIM Daten versionsübergreifend übernehmen kann, ohne an fachlicher Verlässlichkeit zu verlieren. Es dient als Grundlage für die weitere Integration von XFIM in SCOPELAND und für die Abstimmung mit FITKO und KoSIT.

1 Ausgangslage

Ein FIM-Datenschema trägt eine Versionsangabe, ebenso jede darin enthaltene Datenfeldgruppe und jedes Datenfeld. Bei Auswahlfeldern kommt die Version der referenzierten Codeliste im XRepository hinzu. Eine neue Version eines Datenschemas entsteht, sobald sich eines dieser Elemente ändert. Ob die Änderung die Bedeutung der Daten berührt oder nur redaktionell ist, lässt sich an der Schemaversion allein nicht ablesen.

Viele Spezifikationen der öffentlichen Verwaltung sind weder abwärts noch aufwärts verträglich. Eine Datei, die gegen Version 1.2 gültig ist, gilt gegenüber Version 1.0 oder 1.3 formal als ungültig. In der Praxis führt das zur Ablehnung der gesamten Datei, zu manueller Nacherfassung oder zu individuellen Umsetzungsprogrammen zwischen einzelnen Verfahren.

Die tatsächlichen Unterschiede zwischen aufeinanderfolgenden Versionen sind dabei meist gering. Eine Auswertung der Felddaten aus dem FIM-Baukasten zeigt das deutlich. Von 449 Datenschemata liegen 77 in mehreren Versionen vor, bis zu fünf je Datenschema. Über alle 97 Paare aufeinanderfolgender Versionen hinweg sind rund 90 Prozent der Datenfelder der älteren Version in der neueren nach Kennung, Anzahl, Datentyp und Wertebeschränkung unverändert. Etwa 5 Prozent haben sich geändert, etwa 5 Prozent sind entfallen. Selbst ein Sprung der Hauptversion bedeutet nicht

zwingend eine fachliche Änderung. Das Datenschema S05000001287 enthält in Version 1.0.0 und 2.0.0 dieselben 112 Datenfelder mit unveränderten Eigenschaften.

Vergleich aufeinanderfolgender Versionen	Datenfelder	Anteil
unverändert	5.273	89,8 %
geändert (Anzahl, Datentyp oder Wertebeschränkung)	306	5,2 %
in der neueren Version entfallen	296	5,0 %
in der neueren Version hinzugekommen	569	

Grundlage sind die Felddaten aus scpl_fimdf20_fields, verglichen je Datenschema nach Feldkennung. Feldversionen und Gruppenzugehörigkeit sind darin nicht enthalten und flossen deshalb nicht ein.

2 Grundsatz

XFIM folgt dem Grundsatz, genau bei der Identität der Daten und nachsichtig bei ihrer Vollständigkeit zu sein. Genau heißt, dass jeder übernommene Wert eindeutig dem richtigen Datenfeld an der richtigen Stelle in der richtigen Bedeutung zugeordnet wird. Nachsichtig heißt, dass eine Datei nicht deshalb abgelehnt wird, weil sie Felder enthält, die die eigene Version nicht kennt, oder weil Felder fehlen, die die eigene Version verlangt.

In der Softwarearchitektur ist dieses Vorgehen als tolerant reader bekannt. Der Empfänger übernimmt, was er sicher versteht, legt den Rest unverändert ab und macht die Lücken sichtbar. Eine Teilübernahme mit dokumentierten Lücken ist für die Sachbearbeitung wertvoller als eine abgelehnte Datei, solange kein Wert falsch zugeordnet wird. Die folgenden Abschnitte beschreiben die vier Regeln, die das sicherstellen.

3 Zuordnung über den vollständigen Pfad

Die Feldkennungen im FIM-Baukasten sind über Schemaversionen hinweg stabil. F60000000227 bezeichnet in jeder Version dasselbe Datenfeld „Familienname“. Das macht FIM für eine versionsübergreifende Übernahme besonders geeignet, weil ein natürlicher Schlüssel für die Zuordnung vorhanden ist.

Die Feldkennung allein reicht jedoch nicht aus. Dasselbe Datenfeld kommt in einem Datenschema häufig mehrfach vor, etwa als Familienname der antragstellenden Person und als Familienname des gesetzlichen Vertreters. Eine Zuordnung nur über die Feldkennung könnte den Namen des Vertreters der antragstellenden Person zuweisen. Die Zuordnung erfolgt deshalb über den vollständigen Pfad aus Datenfeldgruppen und Datenfeld. Das ist dieselbe Kette von FIM-Kennungen, die auch die Namen der normalisierten Tabellen und der XFIM-Container bildet.

```
G60000000224/G60000000093/G60000000091/F60000000261  (natürliche Person)
G60000000223/G60000000093/G60000000091/F60000000261  (juristische Person)
```

Das Beispiel aus dem Datenschema S00000000362 zeigt, warum der Pfad entscheidet. Das Datenfeld „Staat“ steckt in der Gruppe „Anschrift Ausland“, die über „Anschrift Inland oder Ausland mit Frage“ sowohl beim gesetzlichen Vertreter als natürliche Person als auch beim gesetzlichen Vertreter als juristische Person vorkommt. Nur der vollständige Pfad unterscheidet die beiden Werte.

Ändert sich zwischen zwei Versionen die Gruppenstruktur, wird eine Abbildungstabelle zwischen den Pfaden der Versionen benötigt. Sie lässt sich weitgehend automatisch durch den Vergleich der Pfade erzeugen. Von Hand gepflegt werden nur Fälle, in denen eine Gruppe umbenannt, aufgeteilt oder zusammengelegt wurde. Die Abbildungstabelle ist einmal je Versionspaar zu erstellen und gilt danach für alle Austauschdateien dieses Paares.

4 Verträglichkeitsprüfung je Datenfeld

Gleiche Feldkennung und gleicher Pfad bedeuten noch nicht gleiche Bedeutung. Jedes Datenfeld hat im FIM-Baukasten eine eigene Version. Ändert sich das Datenfeld selbst, kann sich seine Bedeutung verschoben haben, ebenso der Datentyp, die zulässige Länge, das Eingabemuster oder bei Auswahlfeldern die Codelistenversion. Jeder zu übernehmende Wert wird deshalb vor der Übernahme einer von vier Klassen zugeordnet.

Klasse	Merkmal	Behandlung
identisch	gleiche Feldversion, bei Auswahlfeldern gleiche Codeliste in gleicher Version	ohne Rückfrage übernehmen
verträglich	andere Feldversion, aber nur erweiternde Änderung, etwa größere zulässige Länge, weiterer Wertebereich, höhere Obergrenze der Anzahl	ohne Rückfrage übernehmen, im Bericht vermerken
unsicher	andere Feldversion mit einschränkender oder unbekannter Änderung, andere Codelistenversion, anderer Datentyp, kürzere zulässige Länge, niedrigere Obergrenze der Anzahl	anzeigen, Übernahme nur nach Bestätigung durch die Sachbearbeitung
nicht vorhanden	Pfad existiert in der Zielversion nicht	nicht übernehmen, im Original erhalten, im Bericht vermerken

Werte werden nie stillschweigend verändert. Ein Text, der die zulässige Länge der Zielversion überschreitet, wird nicht gekürzt, sondern als unsicher gekennzeichnet. Ein Code, der in der Codelistenversion der Zielversion nicht vorkommt, wird nicht ersetzt, sondern angezeigt. Mehrfachangaben, die die Obergrenze der Zielversion überschreiten, bleiben vollständig im Original erhalten.

5 Führender Datensatz und abgeleitete Fassungen

Werden Daten parallel in mehreren Versionen eines Datenschemas gehalten und unabhängig voneinander bearbeitet, entstehen schnell widersprüchliche Bestände. Deshalb gilt die Regel des führenden Datensatzes. Die empfangene Fassung wird mitsamt der unveränderten Original-Datei aufbewahrt und als führend gekennzeichnet. Fassungen in anderen Versionen desselben Datenschemas gelten als abgeleitet. Sie werden bei jeder Änderung des führenden Datensatzes neu erzeugt und nicht eigenständig bearbeitet.

SCOPELAND verwendet diesen Mechanismus bereits für Datenschemata, die verschiedene Tätigkeiten derselben Leistung beschreiben, etwa die An- und die Abmeldung zur Hundesteuer. Gleichartige Angaben werden dort anhand übereinstimmender Gruppen- und Feldmuster abgeglichen. Die versionsübergreifende Übernahme ist ein weiterer Anwendungsfall desselben Abgleichs.

6 Lückenbericht und Bearbeitungsstatus

Für die Aktenmäßigkeit muss nachvollziehbar bleiben, was übernommen wurde und was nicht. Jeder Import erzeugt deshalb einen Lückenbericht. Er weist für jedes Datenfeld die Klasse aus der Verträglichkeitsprüfung aus, führt die nicht übernommenen Angaben auf und benennt die Pflichtfelder der Zielversion, die leer geblieben sind.

Ein unvollständiger Datensatz wird nicht abgelehnt. Er erhält einen Bearbeitungsstatus, etwa „unvollständig, Nacherfassung erforderlich“ oder „unsichere Werte, Bestätigung erforderlich“. Die Datenaustauschtabellen in SCOPELAND, in denen die Sachbearbeitung schon heute über die Übernahme importierter Daten entscheidet, sind der Ort für diese Prüfung.

7 Übertragung auf XFIM

7.1 Versionsangaben als Bestandteil jeder XFIM-Datei

Die Verträglichkeitsprüfung aus Abschnitt 4 setzt voraus, dass der Empfänger die Version jedes übermittelten Datenfeldes kennt. Bisher enthält XFIM nur die Version des Datenschemas im Namen des Wurzelelements. Ohne Feldversion könnte der Empfänger die Klassen nur raten. XFIM wird deshalb um Versionsangaben auf jeder Ebene erweitert.

Jede Datenfeldgruppe und jedes Datenfeld erhält das Attribut `version` mit der Version aus dem FIM-Baukasten. Auswahlfelder erhalten zusätzlich die Kennung und Version der verwendeten Codeliste. Für diese beiden Angaben werden die Attributnamen `listURI` und `listVersionID` verwendet, wie sie in den XÖV-Standards für Codes üblich sind. Das Wurzelelement trägt die Schemaversion weiterhin im Namen und zusätzlich als Attribut.

```
<S00000000362.1.0.0 version="1.0.0">
  ...
  <G60000000093 version="3.1.0">
    <!-- Abfrage Anschrift Inland oder Ausland -->
    <F60000000263 version="2.0.0">002</F60000000263>
    <G60000000091 version="3.0.0">
      <!-- Staat -->
      <F60000000261 version="2.1.0"
          listURI="urn:de:bund:destatis:bevoelkerungsstatistik:schluessel:staat"
          listVersionID="2023-02-24">...</F60000000261>
      ...
    </G60000000091>
  </G60000000093>
</S00000000362.1.0.0>
```

Die Kennungen und Versionen im Beispiel stammen aus dem Datenschema S00000000362 in Version 1.0.0. Das Auswahlfeld F60000000263 verwendet eine Werteliste im Datenschema selbst, das Feld F60000000261 eine Codeliste aus dem XRepository.

Die Versionsangaben sind Pflichtattribute. In den Schemadefinitionen werden sie mit einem Muster für Versionsnummern geprüft. Bei Auswahlfeldern mit Werteliste im Datenschema selbst entfallen `listURI` und `listVersionID`, die zulässigen Werte stehen dann als Aufzählung in der Schemadefinition.

7.2 Zwei Prüfstufen

XFIM unterscheidet zwischen Versand und Empfang. Beim Versand wird streng gegen die Schemadefinition der eigenen Version geprüft. Nur vollständig gültige Dateien verlassen das Verfahren, damit jeder Empfänger sich auf korrekte Versionsangaben und Werte verlassen kann.

Beim Empfang wird **tolerant** gelesen. Stimmt die Version der eingehenden Datei mit der eigenen überein, gilt die strenge Prüfung auch hier. Weicht sie ab, führt eine Abweichung von der eigenen Schemadefinition nicht zur Ablehnung. Sie fließt stattdessen in die Verträglichkeitsprüfung und den Lückenbericht ein. Abgelehnt wird eine Datei nur, wenn sie kein wohlgeformtes XML ist, wenn das Datenschema nicht erkannt wird oder wenn die Versionsangaben fehlen.

7.3 Ablauf beim Empfang

Der Empfang einer XFIM-Datei läuft in sieben Schritten ab. Zuerst wird die Datei unverändert archiviert und formal auf Wohlgeformtheit, erkanntes Datenschema und vorhandene Versionsangaben geprüft. Stimmt die Version mit der eigenen überein, folgt die strenge Prüfung gegen die eigene Schemadefinition. Andernfalls werden die Pfade über die Abbildungstabelle des Versionspaares zugeordnet. Danach erhält jeder Wert seine Klasse aus der Verträglichkeitsprüfung. Die identischen, verträglichen und unsicheren Werte werden in die Datenaustauschtabellen übernommen, und der Lückenbericht wird erzeugt. Anschließend entscheidet die Sachbearbeitung über die unsicheren Werte und die Übernahme in den Vorgang. Zuletzt werden die abgeleiteten Fassungen in den übrigen Versionen des Datenschemas aktualisiert.

8 Bedeutung für NOOTS und XNachweis

Im Nachweisabruf über NOOTS treffen Behörden mit unterschiedlichen Ständen ihrer Fachverfahren aufeinander. Ein Nachweis, der als XFIM-Datei in der Nutzlast von XNachweis übermittelt wird, bleibt mit dem hier beschriebenen Vorgehen auch dann verwertbar, wenn der Datenabrufende eine andere Version des Datenschemas verwendet. Die Versionsangaben je Datenfeld ergänzen zudem die Verweise des Nachweiskatalogs auf FIM-Kennungen um die fehlende Angabe, welche Fassung eines Datenfeldes tatsächlich übermittelt wurde.

Das Vorgehen schwächt keinen bestehenden Standard ab. Es setzt dort an, wo die Versionspflege in der Praxis zu Brüchen führt, und hält die Strenge beim Versand unverändert aufrecht. Das ist ein tragfähiger Gesprächsansatz gegenüber FITKO und KoSIT.