



Stefan Neubert, Septembre 2023

Format XML SpiGes 1.3

Description du fichier XML pour l'importation des données dans la plateforme SpiGes

À propos de ce document

Nom du projet : NaDB - SpiGes	
N° MJP : 20-102	
Nom du document : Format XML SpiGes 1.3 Description du fichier XML pour l'importation des données dans la plateforme SpiGes	
Auteur(e) : Stefan Neubert, Edith Salgado, Lionel Perret	Date : 07.09.2023
Phase de projet : Relevé	N° d'enregistrement : 622.10

version :	Date	But :
1.0	31.03.2023	Première publication
1.2	12.05.2023	Adaptations à la structure XML (déplacement des attributs de l'élément Cas / « Fall »)
1.3	07.09.2023	Extension pour les demandes cantonales et correction d'erreurs

Distribution : Partenaires externes (entreprises et cantons via le site Internet)

Modifications par version et rétrocompatibilité : Version 1.3 Points sans rétrocompatibilité (des modifications peuvent être nécessaires) : - Pour le traitement automatisé, la version XSD a été ajoutée comme attribut propre à l'élément racine « unternehmen ». Il faut indiquer ici le numéro de version, c'est-à-dire actuellement "1.3". - Pour les variables avec des nombres décimaux, le nombre de chiffres après la virgule a été défini plus précisément dans la liste des variables. Dans le XSD, les décimales ont également été adaptées et les types de données des variables de facturation ont été adaptés aux directives du Forum Datenaustausch. - Pour les attributs « ktr_31_6b2 » et « ktr_31_6b2_ank », il manquait un trait de soulignement (_) dans le nom par rapport aux autres attributs. Ce point a été corrigé dans le XSD et dans la liste des variables afin de respecter une systématique uniforme. - Pour les variables « ktr_31_1-5 » et « ktr_31_1-5_ank », un trait d'union apparaissait dans le nom de la variable. Comme cela pouvait enfreindre certaines conventions de dénomination, le trait d'union a été remplacé par un soulignement. - La faute d'orthographe dans l'attribut « aufnemmengewicht » et dans le nom de variable du même nom a été corrigée. Les deux s'appellent désormais « aufnahmegewicht ». - Pour certaines variables, il y avait une incohérence concernant la validité de la valeur 0 entre le XSD et la liste des variables. La définition XSD a été adaptée à la liste de variables. Variables concernées : « geburtsgewicht », « laenge », « aufnahmegewicht », « psy_austritt_behandlung ». - L'attribut « behandlung_id » apparaît dans l'élément « Operierende ». Celui-ci n'est toutefois pas nécessaire en raison de la relation hiérarchique dans le XML. Il a donc été supprimé. - L'élément « burnr_gesv » est obligatoirement requis par certains partenaires. Il est donc obligatoire et pas seulement facultatif. Points avec compatibilité ascendante (aucune modification nécessaire) : - L'élément « Kantonsdaten » peut désormais être subordonné non seulement à l'élément « Fall », mais aussi à l'élément « Standort » et à l'élément « Unternehmen ». Ainsi, les cantons peuvent également saisir des données supplémentaires à ce niveau.
--

- j. La variable « alter_U1 » a été ajoutée comme attribut de l'élément « Administratives ». Cette indication est facultative et sert à ce que le fichier de données SpiGes puisse être regroupé selon SwissDRG, TARPSY, ST-Reha et SPLG.
- k. Les erreurs lors du contrôle de validité des attributs « eintrittsart » et « psy_eintritt_aufenthalt » ont été corrigées.
- l. Le type de données de la variable « medi_dose » a été modifié d'entier à décimal, comme cela pouvait déjà être fourni dans le MS.
- m. Les valeurs autorisées de la variable « psy_einweisung_instanz » ont été corrigées dans le XSD et alignées sur la liste des variables.

Version 1.2

- a. Les attributs de l'élément cas (« Fall ») sont désormais représentés dans des éléments propres situés sous l'élément cas (« Fall »). Les nouveaux éléments s'appellent administratives (« Administratives »), psychiatrie (« Psychiatrie ») et nouveau-né (« Neugeborene »).
- b. La structure du fichier des identifiants change de manière analogue. L'élément "Identifiant de personnes" est désormais représenté sous l'élément de cas (« Fall ») nouvellement créé.
- c. L'attribut « medi_id » a été ajouté à l'élément médicaments « Medikamente ».
- d. L'attribut « rech_id » a été ajouté à l'élément factures « Rechnungen ».
- e. L'attribut « burnr_gesv » a été ajouté à l'élément administratifs « Administratives ».
- f. La longueur de l'attribut "medi_einheit" a été augmentée de 3 à 4.

Compatibilité ascendante : les points a à d doivent impérativement être mis en œuvre. Il n'y a pas de compatibilité ascendante en ce qui concerne ces derniers.

Table des matières

1	Objectif du document	5
2	Format	5
2.1	Format XML	5
2.2	Livraison des données en deux fichiers.....	5
2.3	Structure du fichier de données.....	5
2.4	Éléments facultatifs et attributs	8
2.5	Élément centre de coûts	8
2.6	Structure du fichier d'identifiants	8
2.7	Données cantonales supplémentaires	9
2.8	Encodage.....	9
2.9	«Namespaces» (Espaces de nommage).....	9
2.10	Étendue et délimitation de la fourniture de données.....	10
2.11	Adaptations du format XML.....	10
3	Transmission des fichiers	11
3.1	Désignation	11
3.2	Taille des fichiers et compression	11
3.3	Livraisons partielles	11
3.4	Contrôle du format.....	12

1 Objectif du document

Le document décrit le format XML utilisé pour transmettre les données SpiGes des hôpitaux à la plateforme de relevé. Le document décrit le contenu, la structure et la conception technique du format. D'autres informations sont disponibles sur la [page d'accueil SpiGes](#), notamment le concept détaillé (indications sur le projet et le déroulement du relevé, concepts fondamentaux), la liste des variables (indications sur les différentes variables relevées), un exemple de fichier XML, ainsi que le schéma XSD (définition et contrôle du document XML).

2 Format

2.1 Format XML

Les informations demandées sont transmises sous forme de fichier XML (Extensible Markup Language).

Contrairement aux formats précédents, le format n'est plus plat mais structuré, ce qui facilite les livraisons partielles et les adaptations. Le fichier XML est décrit dans un schéma XSD. Ce schéma est également disponible sous une forme structurée et peut être utilisé pour vérifier le format des fichiers XML.

2.2 Livraison des données en deux fichiers

La livraison des données à la plateforme SpiGes comprend deux fichiers : un fichier de données et un fichier d'identifiants. Le premier fichier contient toutes les données "de contenu" de l'enquête. Le fichier des identifiants comprend les identifiants des personnes, c'est-à-dire concrètement le numéro AVS, la date de naissance ainsi que l'identifiant du cas.

La séparation des deux fichiers est l'une des nombreuses mesures visant à garantir la sécurité des données. Ainsi, les données personnelles sensibles ne sont pas transmises avec les données d'identification (n° AVS). En cas de livraison de corrections, seul le fichier erroné doit être livré, de sorte que le fichier d'identification ne doit idéalement être livré qu'une seule fois et que toutes les autres livraisons de corrections ne comprennent que le fichier de données. Ainsi, le risque pour la sécurité des données peut être massivement réduit tout au long du processus.

2.3 Structure du fichier de données

Le fichier de données est divisé hiérarchiquement en trois niveaux : entreprise (« Unternehmen »), site (« Standort ») et cas (« Fall »). Pour les cas, il est possible d'indiquer les éléments suivants : administratives (« Administrative »), nouveau-né (« Neugeborene »), psychiatrie (« Psychiatrie »), diagnostic (« Diagnose »), traitement (« Behandlung »), médicament (« Medikament »), facture (« Rechnung »), centre de coût pour le cas (« Kostentraegerfall »), mouvement du patient (« Patientenbewegung ») ainsi que des données cantonales optionnelles (« Kantonsdaten »). Pour les traitements, il est en outre possible d'indiquer les opérateurs (« Operierende ») comme sous-élément. A l'exception des identifiants, des objets de coûts liés au site (« Standort ») à l'entreprise (« Unternehmen ») et données cantonales, tous les attributs se situent en dessous de l'élément cas (« Fall »).

Les éléments entreprise, site et cas n'ont pas d'autres attributs que leur identifiant respectif (ent_id, burnr). Cette structure permet de transmettre différentes parties de l'enquête (p. ex. données sur les objets de coûts) dans un fichier XML séparé sous forme de livraison partielle.

Chaque cas (« Fall ») est clairement attribué à un site (« Standort »), à savoir le site principal. Voir à ce sujet la spécification de la variable « burnr ».

La structure schématique du fichier de données est représentée ci-dessous (le fichier XML n'est pas traduit, donc l'illustration suivante est en allemand, traduction de termes, voir ci-dessus) . Pour des raisons de clarté, la plupart des attributs sont masqués. Les différentes variables (Attributs), y compris l'élément XML supérieur, figurent dans la liste des variables du SpiGes.

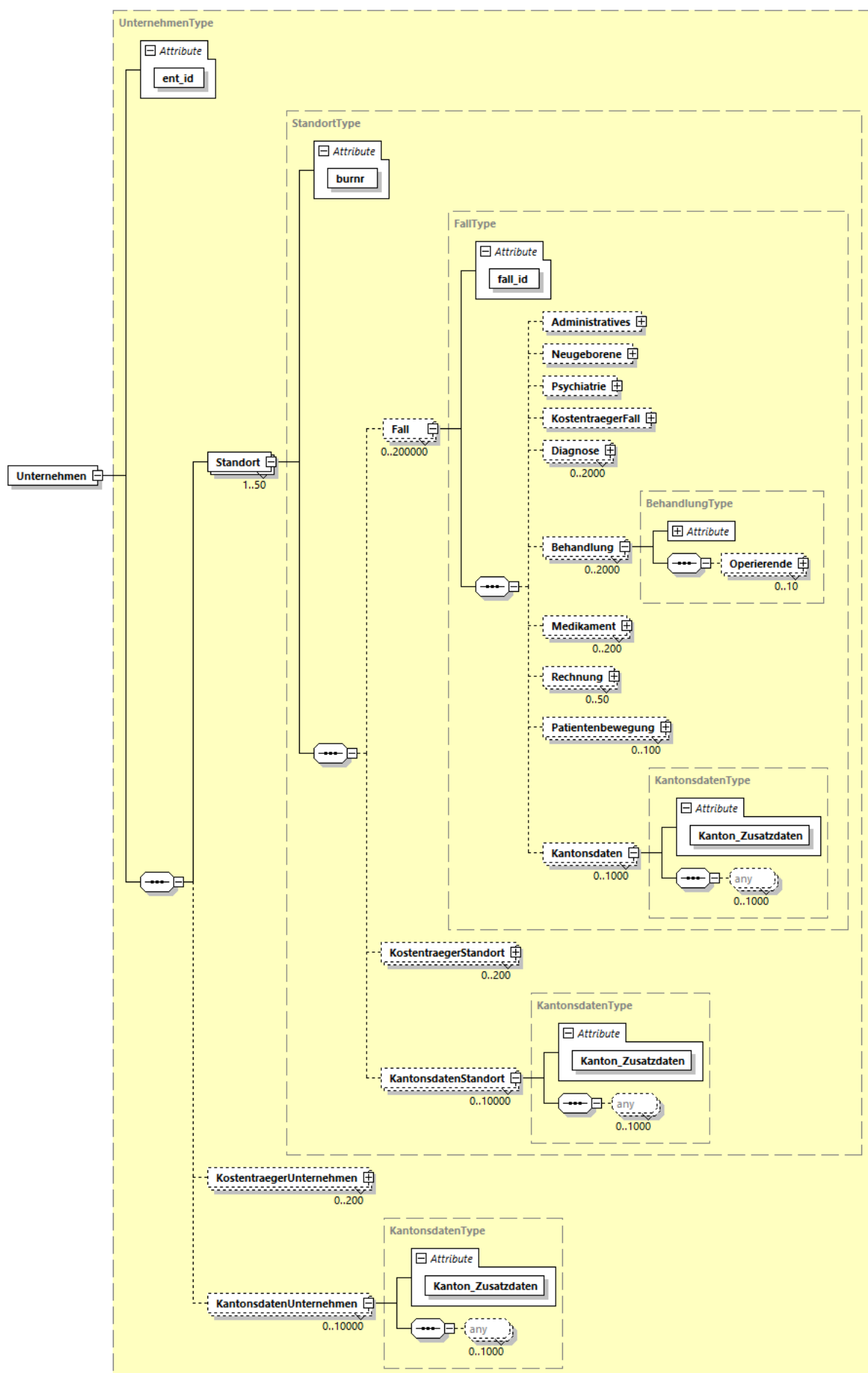


Illustration 1: Structure hiérarchique du fichier de données

2.4 Éléments facultatifs et attributs

Les éléments optionnels sont représentés par des lignes pointillées. Les éléments optionnels sont techniquement facultatifs, c'est-à-dire qu'un fichier est reconnu comme valide même si les éléments correspondants manquent. Cela ne signifie toutefois pas qu'il est correct du point de vue du contenu de ne pas fournir ces éléments du tout. Les éléments sont optionnels pour différentes raisons :

- *Les cas* devraient en principe se trouver dans chaque enregistrement. Il se peut toutefois que seuls des objets de coûts (PIG ou autres) soient attribués à un site administratif, mais pas de cas. Dans ce cas, aucun cas ne devrait être indiqué pour un site.
- Les informations *administratives, nouveau-né, psychiatrie, les diagnostics, les traitements, les médicaments, les factures, les centres de coût pour le et les mouvements de patients* peuvent être livrés séparément et sont donc facultatifs. Il est ainsi possible, par exemple, de faire contrôler au préalable un ensemble de données sans informations sur les centres de coût du cas.
- *Les opérateurs* sont facultatifs, car les variables associées sont facultatives et ne doivent être saisies que selon les instructions du canton.
- *Les données cantonales* sont facultatives, car elles ne doivent être indiquées que si le canton a donné des instructions et des spécifications en ce sens (voir section 2.7 Données cantonales supplémentaires).
- *Les centres de coût entreprise (« KostenträgerUnternehmen ») et les centre de coût site (« KostenträgerStandort »)* sont facultatifs, étant donné que les données relatives aux centres de coût peuvent être fournies à différents niveaux (cf. section 2.5 Élément centre de coûts) et parce que les données relatives aux centres de coûts peuvent être fournies séparément.

Les attributs optionaux sont des variables qui ne doivent pas obligatoirement être présentes pour que le document XML soit reconnu comme un fichier valide. Cela concerne par exemple les variables qui ne doivent être indiquées que pour les sorties ou encore certaines variables, comme le pays de résidence, qui ne doit être indiqué que si le canton de résidence n'a pas été rempli parce que la personne est domiciliée à l'étranger.

2.5 Élément centre de coûts

L'élément centre de coûts peut être indiqué à différents niveaux. Il peut d'une part être rattaché à l'élément cas (objets de coûts liés au cas). Cela correspond à l'ancienne statistique des coûts par cas ou au fichier des coûts par cas de SwissDRG SA. D'autre part, les objets de coûts peuvent également être attribués au site ou à l'entreprise. Cela n'est toutefois autorisé que pour les objets de coûts non liés à un cas, c'est-à-dire par exemple les établissements secondaires, les PIG, les mandats ou les informations sur les objets de coûts ambulatoires saisies de manière agrégée.

Il faut veiller à ce que les objets de coûts au niveau de l'entreprise ne soient indiqués que si l'ITAR_K® doit également être établi au niveau de l'entreprise. Si plusieurs ITAR_K® doivent être créés pour une entreprise (c'est-à-dire au niveau d'un ou de plusieurs sites), les informations sur les objets de coûts doivent toujours être disponibles au niveau du site ou du cas. En d'autres termes, SpiGes ne peut pas répartir des objets de coûts existant au niveau de l'entreprise sur un ou plusieurs sites afin de reproduire un ITAR_K® pour ceux-ci.

2.6 Structure du fichier d'identifiants

La structure du fichier d'identifiants est la même aux niveaux supérieurs que la structure du fichier de données : les entreprises et les sites sont identiques. Au lieu de l'élément cas, ce sont les éléments identifiants des patients (« Personenidentifikatoren ») qui sont indiqués au niveau du site. Un élément d'identifiant du patient doit être indiqué pour chaque cas.

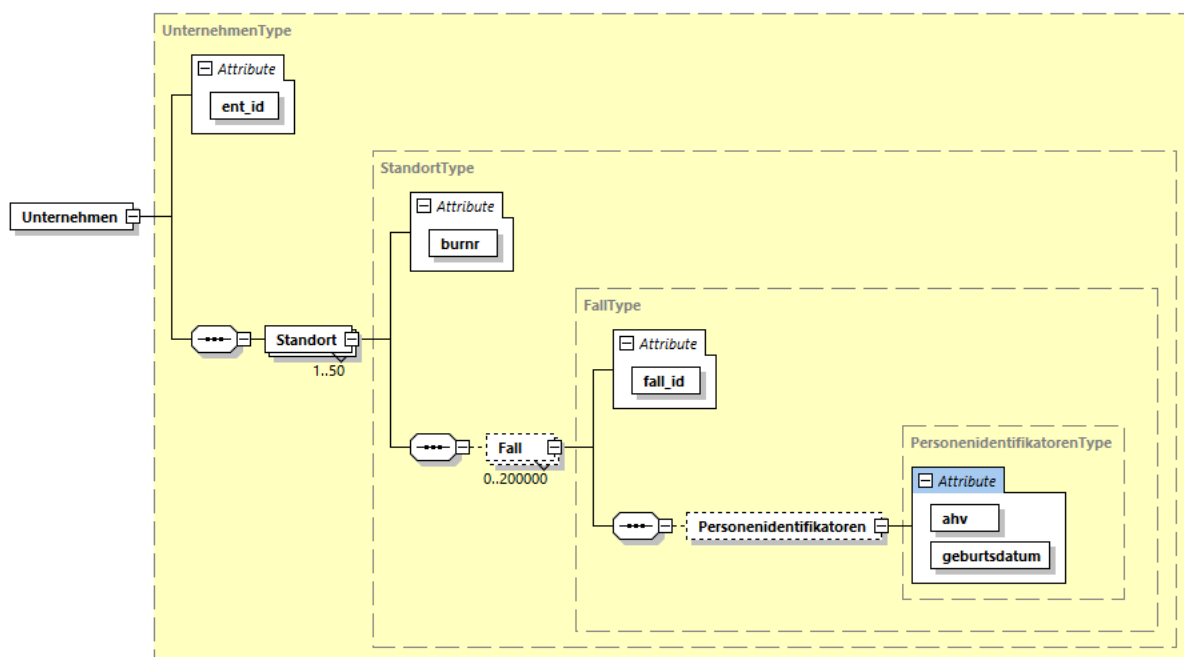


Illustration 2: Structure hiérarchique du fichier d'identifiants

2.7 Données cantonales supplémentaires

Les éléments optionnels "Données cantonales" peuvent être attribués à l'élément "cas", à l'élément « site » ou à l'élément « Unternehmen » et peuvent être aménagés à volonté avec des sous-éléments selon les directives du canton. L'OFS recevra ces données et les transmettra au canton sans les examiner. Aucun contrôle de cas individuels ou d'indicateurs n'est prévu sur ces données. Tous les données cantonaleléments doivent comporter l'attribut "Kanton_Zusatzdaten", dans lequel est indiqué le canton qui fournit les données supplémentaires. Ainsi, il est également possible qu'une entreprise hospitalière fournisse des données supplémentaires pour plusieurs cantons en même temps. Les cantons ne reçoivent alors que leurs propres données supplémentaires.

Pour la collecte de leurs propres données cantonales supplémentaires, les cantons doivent communiquer les spécifications directement aux entreprises hospitalières concernées. L'OFS doit également être informé. Outre une description des variables, il faut un fichier XSD décrivant la structure de l'élément XML « Kantonsdaten ». Lors de la conception des données supplémentaires cantonales, il faut tenir compte des directives de l'association eCH concernant la conception de l'échange électronique de données (cf. [eCH-0018](#)).

2.8 Encodage

Les fichiers XML doivent être sauvegardés avec l'encodage UTF-8. Cela permet notamment de s'assurer que les caractères spéciaux dans les champs de texte libre sont correctement restitués.

2.9 «Namespaces» (Espaces de nommage)

Les « Namespaces » permettent de garantir que les définitions dans le fichier XML peuvent être attribuées de manière univoque. Il s'agit quasiment du niveau supérieur de l'arborescence hiérarchique du fichier XML. Les namespaces sont indiqués sous forme d'URL (adresses internes). Le « Namespace » du fichier XML de SpiGes est "http://www.bfs.admin.ch/xmlns/gvs/spiges-data/1.0" pour la version 1.0 du XML de données et sera adapté en conséquence pour les versions suivantes. Il ne s'agit pas d'une adresse Internet fonctionnelle. La définition sous cette forme permet uniquement de garantir l'unicité au niveau mondial. La définition générale du format XML est utilisée comme autre « Namespace » ("http://www.w3.org/2001/XMLSchema"). Les cantons qui définissent des données supplémentaires

doivent également indiquer un « Namespace » dans leur fichier XSD.

2.10 Étendue et délimitation de la fourniture de données

Une livraison de données SpiGes comprend tous les cas d'hospitalisation de la période de relevé ainsi que d'autres informations ponctuelles (p. ex. sur les centres de coût). Les définitions de la population de base et des cas ainsi que d'autres concepts centraux sont fixés dans le concept détaillé et se basent pour l'essentiel sur les relevés précédents (Statistique « MS », Statistique « Fallkostenstatistik/SwissDRG-Kostendaten »).

2.11 Adaptations du format XML

En principe, les formats XML visent une compatibilité descendante. C'est un avantage pertinent par rapport au format à caractères séparés utilisé dans la MS actuelle. Il est ainsi possible de faire des extensions sans que les formats existants ne fonctionnent plus. En cas de nouveautés, on communique à chaque fois de manière détaillée quelles modifications doivent être effectuées obligatoirement et lesquelles sont optionnelles.

3 Transmission des fichiers

3.1 Désignation

En principe, le nom du fichier peut être choisi librement. Seule l'extension en ".xml" doit être respectée. Le contenu et le format des fichiers sont contrôlés lors du téléchargement. Les fichiers de données sont renommés pour la gestion interne et le téléchargement avec l'extension "..._data.xml", les fichiers d'identifiants pour l'action interne avec l'extension "..._ids.xml".

3.2 Taille des fichiers et compression

La taille des fichiers XML peut être nettement supérieure à 100 Mo. Pour garantir une transmission efficace, les fichiers peuvent être comprimés avant d'être transmis. Pour ce faire, il convient d'utiliser un algorithme de compression « deflate » et de nommer le fichier avec l'extension ".zip". Sur la plateforme, le nombre d'éléments contenus dans les fichiers ZIP n'est pas affiché.

3.3 Livraisons partielles

Comme indiqué dans la section 2.4 "Éléments facultatifs" ce format permet également d'effectuer des livraisons partielles. Ceci est par exemple utile lorsque toutes les données ne sont pas encore prêtes, mais que les données existantes doivent tout de même déjà être vérifiées.

Les livraisons partielles peuvent être effectuées par combinaison d'élément et de site, par exemple pour les diagnostics du site A et les *centres de coûts site* (« KostenträgerStandort ») du site A ainsi que les *centres de coûts site* (« KostenträgerStandort ») du site B. Les éléments sont les informations administratives, les nouveau-nés, la psychiatrie, les diagnostics, les traitements, les médicaments, les factures, les centres de coût pour le cas du site A et les mouvements de patients ainsi que les centres de coûts entreprise et les centres de coûts site. Les livraisons partielles ne peuvent toutefois pas être réparties de manière plus fine.

Lors de chaque téléchargement, le système vérifie s'il existe déjà des données pour cette combinaison de lieu et d'élément. Si des données existent déjà, elles sont remplacées. Cela signifie que seules les données du dernier téléchargement sont encore valables pour cet élément sur le site. Inversement, si la livraison ne contient pas de données pour une combinaison de lieu et d'élément, les dernières données livrées restent valables pour cette combinaison d'élément et de lieu.

En d'autres termes, cela signifie qu'il n'est pas possible de fournir des données complémentaires pour une telle combinaison, par exemple les informations administratives pour une première partie des cas du site X dans une livraison et la deuxième partie des informations administratives des cas du site X dans une livraison suivante. Dans ce cas, seuls les cas de la deuxième livraison seraient conservés. Les cas de la première livraison seraient perdus.

Le tableau suivant donne un exemple de ce qui se passe avec les livraisons partielles :

Site	Élément	Données valables avant le téléchargement	Téléchargement	Données valables après le téléchargement
71284578	Centre de coûts Site	xxx	zzz	zzz
71284578	Centre de coûts Site	yyy		
71284578	Administratives	aaa	bbb	bbb
71284578	Administratives		ccc	ccc
71284578	Diagnostic	ddd	ddd	ddd
71284578	Diagnostic	eee	eee	eee
71284578	Diagnostic	fff	ggg	ggg

71284578	Centre de coûts Cas		hhh	hhh
71284578	Centre de coûts Cas		iii	iii
88888888	Centre de coûts Site	aaaa	aaaa	aaaa
88888888	Cas	kkk		kkk

Illustration 3: Exemples de livraisons partielles

3.4 Contrôle du format

Lors du téléchargement du fichier sur la plateforme, le format du fichier est contrôlé à l'aide du schéma XSD. Les erreurs sont signalées et doivent être corrigées avant que les éléments erronés puissent être importés et traités. Le contrôle suit la même logique que celle utilisée pour les livraisons partielles (cf. paragraphe précédent 3.3 Livraisons partielles).

Le contrôle XSD peut être effectué dès l'hôpital. Cela permet un contrôle aussi proche que possible de la source. Des contrôles plus poussés (contrôles au cas par cas, contrôles des ratios) sont ensuite effectués sur la plateforme.