

Base de données de RSA : Les choix, l'orientation générale

Catégorie : Base de données de RSA

Publié par Philippe Feigel [[admin](#)] le 19/09/2006

Le but de cette base de données est de faciliter l'utilisation de base de RSA à caractère régionale (ou autre) et l'exploitation autour du PMSI dont des dérivés discutables certes mais réels comme la planification, les parts de marché, l'épidémiologie avec leurs limites. Les difficultés d'utilisation de ces RSA sont la multiplicité des formats, des nomenclatures difficiles à obtenir ou à maintenir, des outils d'exploitations inexistantes et l'absurdité de recréer dans chaque établissement cet ensemble au dépens du temps d'analyse.

Plusieurs outils permettent dès à présent d'interroger des bases nationales comme par exemple les sites de [Parhtage](#), [ATIH](#), [BDHF](#) mais sont conçus autour de requête type répondant plus à des notions de tableaux de bords qu'à des études comme : quelle est la part de médecine gériatrique de ce secteur ou quelle répartition des séances de chimio d'une catégorie de cancer...

A l'initiative des délégations régionales des fédérations d'hospitalisations (FHF, FHP, FEHAP) avec le soutien et le financement de l'ARH des Pays de Loire ce travail a été initié, dans une optique collaborative et une licence libre. Il vous est demandé de citer les auteurs en cas de publication ou de travaux et de mentionner le site et la mission, mais plus encore de participer à l'évolution en particulier les nomenclatures manquantes ou les rapports de bugs.

Les données sont à obtenir auprès de votre ARH pour les données régionales, l'ATIH pour des données nationales, mais la conception de la base permet un fonctionnement par collecte de fichiers de RSA comme par exemple auprès d'établissements (réseaux, fédérations ...).

Les choix :

Absence de perte d'information :

L'ensemble des données des RSA est conservé, sans aucune perte des données originales. Le type des données est respecté (numérique, chaîne, logique) suivant le sens qu'ils ont par exemple : durée = numérique, Code GHM = Chaîne. Le type logique n'est que rarement utilisé et uniquement si il n'y a pas de possibilité de valeur nulle dans le format d'origine. Les données ambiguës comme le Code GHS sont en format chaîne tant que le format permet un classement ascendant logique. Il n'y a jamais de transformation de données. La conversion de champs en numérique du fichier de données est assurée par le moteur de base de données (normalement sans surprise pour ces types simples de données).

Formats Pivots : Les données de RSA sont organisées dans des formats pivots autant pour la partie fixe que les parties variables comme les actes. Ce principe permet d'exploiter plusieurs années en une seule requête et réaliser des comparaisons pluri-annuelles.

Il est bien sûr nécessaire de bien connaître le PMSI et ses évolutions dont l'existence de certaines données à telle ou telle année. Les différents formats des RSA et leurs contenus sont rappelés dans l'outil de découpage de RSA et son aide intégrée. Certains champs comme le diagnostic principal, l'âge sont communs à tous les formats, d'autres que dans un format.

En l'absence de nom partagé pour les champs, ceux-ci ont été construits pour être le plus descriptifs. En sus, des commentaires de la structure des tables apportent des renseignements

supplémentaires.

Le format pivot est consultable facilement dans la structure de la base.

Une limite au format pivot est le changement complet de référentiel. Par exemple, bien qu'il ait été possible d'avoir un seul format pour les actes en CDAM et en CCAM, l'exploitation en aurait été trop difficile et aurait impliqué une nomenclature d'acte unique, trop disparate pour une maintenance efficace. Il existe donc deux tables une pour les actes en CDAM, et une table d'acte en CCAM dans un format pivot permettant de gérer les évolutions d'informations.

Formats supportés : Les formats supportés débutent à partir du format de RSA 205. Les versions antérieures sont trop éloignées en contenu avec une faible pertinence de données de plus de 5 ans et des référentiels lointains comme la CIM 9 (HCIMO). En sus, partiellement supporté, les fichiers de RSA déjà découpés pour certaines versions (voir l'aide).

Partie fixe et variable :

Les RSA sont constitués d'une partie fixe, et de parties répétitives variables (Actes, Unités, Diagnostics). Chacune de ces parties constitue une ou des tables.

La partie fixe et les parties répétitives sont liées par une clé unique.

Certaines informations sont dupliquées comme par exemple le diagnostic principal à la fois dans la partie fixe, mais aussi répété dans la table diagnostic permettant ainsi des requêtes rapides sur les diagnostics principaux mais simplifiant les requêtes nécessitant l'analyse des diagnostics quelque soit leurs positions.

N° d'import :

Il est nécessaire de créer une clé unique pour les imports. Le candidat à clé unique serait N°Finess+N°RSA+AnnéeRSA, mais les usages possibles de cette base peuvent rencontrer des situations où les N° de RSA sont réutilisés. Donc à l'usage, il a été créé un N° import. Ce numéro associé au n° rang du RSA dans le fichier compose une clé unique de RSA et à l'avantage de permettre une suppression simple d'import. En attendant de gérer une connexion à la base avec DecoupeRSA, il faut être vigilant, vérifier le N° Import et l'incrémenter manuellement.

Nomenclature :

Les nomenclatures sont prévues pour une correspondance limitant le risque de produit cartésien lors de requêtes. Ce choix est dicté par la facilité d'analyse qu'il implique. Par exemple la nomenclature des GHM comprend la totalité des codes de GHM quelque soit l'année et la fonction de groupage, donc un casemix sur plusieurs années produira une seule ligne par code de GHM et un seul libellé. Toutefois, cela implique l'absence d'historisation des nomenclatures, historisation lourde à gérer, difficile à implémenter pour des requêtes pluri-annuelles.

Des choix implicites : Les libellés parfois d'autres éléments quand ils sont modifiés avec les années sont ceux de la dernière version de la nomenclature. Par exemple la CCAM fournie comprend les versions à partir de la v0bis, mais les hiérarchies sont différentes. Le choix a été de retenir la dernière version, les actes n'existant pas dans la dernière version ont été manuellement reclassés. Autant que possible, les hiérarchies des nomenclatures sont associées comme par exemple CMD et GHM. Les relations entre les tables sont indiquées au moteur de bases.

Aide & Documentation :

L'aide, la documentation et les licences sont systématiquement associées aux téléchargements.

Bien qu'il n'y est pas de garantie de support, vous pouvez nous contacter directement (coordonnées en page d'accueil du site MARTAA), ou nous poster les questions ou bugs, et nous essaierons de vous assister.