

# Vers la fusion d'informations hétérogènes et partielles pour l'aide au codage diagnostique

Laurent Lecornu\*,\*\*\*, Clara Le Guillou\*\*,\*\*\* Frédéric Le Saux\*, Matthieu Hubert\*\*, Julien Montagner\*,\*\*\*, John Puentes\*,\*\*\*, Jean-Michel Cauvin\*\*,\*\*\*

\* Telecom Bretagne , Technopole Brest Iroise, CS 83818  
29238 Brest Cedex 3

Laurent.Lecornu@telecom-bretagne.eu

\*\*CHRU de Brest, DIM, Boulevard Tanguy Prigent  
29609 Brest Cedex

jean-michel.cauvin@chu-brest.fr

\*\*\*Latim, Inserm U650

## 1 Contexte

En effet, depuis la mise en place du Programme de Médicalisation des Système d'Information (PMSI), l'ensemble des informations médicales concernant le patient au cours de son séjour à l'hôpital doit être recueilli sous une forme standardisée et codée. Malheureusement, choisir un code diagnostique est une opération non-intuitive pour le praticien en raison de la quantité de codes prévue par les normes, et de la difficulté d'en établir une classification pertinente. Les erreurs sont fréquentes avec des conséquences graves tant au niveau de l'évaluation des soins de santé qu'au niveau financier. Pour aider le médecin-codage, l'approche choisie consiste à lui présenter une liste de codes diagnostiques adaptée au cas du patient dont il code le séjour, ordonnée par ordre de pertinence, et correspondant aux éléments diagnostiques qu'il a en tête à partir des informations réparties au sein du Système d'Information Hospitalier (SIH) et sur des connaissances externes. Dans ce but, un ensemble de modules spécifiques d'aide au codage a été développé. Chacune d'entre eux analyse une partie spécifique des données du SIH et en extrait une information partielle sous la forme d'une liste ordonnée et contextualisée de codes diagnostiques classés selon une valeur de pertinence qui lui est propre. L'ensemble des listes est alors fusionné pour ne présenter au médecin qu'une seule liste ordonnée. Deux stratégies d'ordonnement de ces valeurs sont explorées (la 1<sup>ère</sup> méthode est dérivée d'une fusion procédurale et la 2<sup>ème</sup> utilise la théorie des possibilités).

## 2 Méthode

Le premier module développée s'appuie sur une étude statistique des codes diagnostiques en fonction de l'âge du patient, de son sexe, de la durée du séjour, de l'unité médicale fréquentée, des actes médicaux effectués et des diagnostics déjà codés (Lecornu (2009b)). Le second s'appuie sur le codage de séjours antérieurs en tenant compte du temps écoulé entre