



**HAL**  
open science

## **Introduction (RSTI 37 : Numéro spécial TSI )**

Nathalie Aussenac-Gilles, Nathalie Bricon-Souf, Lynda Tamine

### ► **To cite this version:**

Nathalie Aussenac-Gilles, Nathalie Bricon-Souf, Lynda Tamine. Introduction (RSTI 37 : Numéro spécial TSI ). Revue des Sciences et Technologies de l'Information - Série TSI : Technique et Science Informatiques, 2018, Information médicale : représentation et traitement, 37 : Numéro spécial TSI - Technique et Science Informatiques (1-6), pp.7-9. <10.3166/tsi37-0006>. <hal-02924569>

**HAL Id: hal-02924569**

**<https://hal.science/hal-02924569v1>**

Submitted on 29 Aug 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## INTRODUCTION

---

Il est indéniable que la santé et le bien-être sont des questions sociétales de premier plan. Ces dernières années, la contribution du numérique dans la résolution de ces questions en France et dans le monde prend de plus en plus une place prépondérante comme en témoigne l'évolution croissante des systèmes d'information en santé. En effet, l'information médicale est devenue de plus en plus un objet d'intérêt majeur pour une large variété d'utilisateurs (patients et leurs familles, médecins, cliniciens, chercheurs, etc.) à des fins d'usages très diversifiés (assistance au diagnostic, recherche clinique, prédiction de pathologie, etc.).

Si certains systèmes (au sein des hôpitaux ou des cabinets, l'imagerie par exemple) atteignent une certaine maturité en France pour un usage ciblé (gestion médico-économique, aide à la prescription, etc.), ils doivent aujourd'hui être capables de s'interfacer avec d'autres systèmes pour différents usages à fort impact sociétal comme l'aide au diagnostic, la recherche clinique ou encore l'assistance à des études épidémiologiques. En outre, l'introduction de nouveaux modes de production et de partage des données (comme les capteurs ou alors les réseaux sociaux) et l'ambition de converger vers une médecine personnalisée font émerger de nouveaux défis scientifiques relatifs à la gestion de données massives (Big data), à l'interopérabilité sémantique entre des données de plus en plus complexes, hétérogènes et de qualité variable, à l'efficacité des modes d'accès à l'information et de découverte de traits saillants. Ces défis nous interrogent sur l'évolution et la place de l'information médicale dans ces applications, tant sur les volets de la représentation, de son traitement et de son exploitation efficace.

Ces questions sont abordées par les communications présentées lors des Symposium sur l'Ingénierie de l'Information Médicale. Faisant suite au 4<sup>e</sup> Symposium sur l'Ingénierie de l'Information Médicale (SIIM 2017) organisé à Toulouse en novembre 2017, ce numéro thématique rend compte des travaux représentatifs de cette évolution. Ce volume comporte cinq articles qui fournissent un aperçu des travaux actuels du domaine concernant l'aide à la décision médicale, la fouille de données et le traitement de corpus médicaux.

L'article de B. Seroussi et ses co-auteurs « *Une plateforme multimodale d'aide à la décision, application à la prise en charge du cancer du sein dans le cadre du projet DESIREE* » présente ainsi un système d'aide à la décision qui combine différentes approches utilisées par les professionnels de santé afin de nourrir leur démarche de décision : la prise en compte des guides de bonnes pratiques, l'utilisation d'expériences acquises au fil de la rencontre de cas complexes, et la comparaison avec des décisions prises pour des cas similaires. Le système proposé met en œuvre ces trois démarches au moyen de méthodes d'aide à la décision complémentaires, proposant respectivement une formalisation des recommandations de bonne pratique, de règles issues des expériences cliniques, et des cas cliniques

rencontrés. Ces méthodes sont interoperantes grâce à un modèle ontologique partagé, et proposent des outils de visualisation adaptés.

L'article de G. Chahbandarian et ses co-auteurs « *Une approche pour la sélection de variables stables : application à l'encodage de diagnostic secondaires* » s'inscrit également dans une perspective d'aide pour les cliniciens. L'exploration des variables contenues dans les bases médico-économiques du PMSI a ainsi l'objectif de leur proposer des guides pour l'encodage des diagnostics secondaires, s'appuyant sur quelques variables pertinentes sélectionnées par le système. Une méthodologie de sélection des variables stables et pertinentes est ainsi proposée pour ce cadre. Le PMSI est une base de données facilement accessible mais ne contenant que peu d'informations médicales. Au-delà de la proposition d'approche pour la sélection des variables, il est intéressant de voir que les résultats obtenus par un système de prédiction s'appuyant sur cette sélection sont très encourageants.

L'article de J. Pinaire et ses co-auteurs « *Nouvelle approche de caractérisation des flux de patients* » utilise également les données du PMSI afin de mieux appréhender les trajectoires hospitalières des patients. Cette connaissance est importante pour optimiser la planification des soins. Le travail est décliné pour des patients hospitalisés pour infarctus du myocarde. Des méthodes de fouille de données portant sur des motifs spatio-temporels sont appliquées, en substituant la proximité spatiale des éléments médicaux par la proximité de leurs pathologies. Des outils de visualisation des trajectoires ainsi obtenues sont proposés. La classification des trajectoires observées permet de définir des profils types de patients en termes d'occurrence et durée d'hospitalisation.

L'exploitation des informations contenues dans les textes présente également un enjeu considérable, compte tenu du volume important de données textuelles utilisées en santé. L'article de C. Dalloux et ses co-auteurs « *Détection de la négation : corpus français et apprentissage supervisé* » répond à un objectif de détection de négation dans les textes. Un corpus contenant des protocoles d'essais cliniques et des cas cliniques est constitué pour disposer de données de référence en Français. Il est ensuite annoté manuellement pour permettre l'apprentissage. Les auteurs présentent les méthodes actuelles utilisées ainsi que leurs propositions s'appuyant sur des réseaux de neurones récurrents, pour la détection des marqueurs de négation ainsi que pour la détection de la portée de ces marqueurs.

Etudier les textes d'utilisateurs de forums de santé permet de comprendre les besoins et la compréhension qu'ont ces utilisateurs de leur santé. L'article de T. Hamon et ses co-auteurs « *Fouille de forums de discussion en santé : identification des caractéristiques des contributeurs* » cherche à caractériser les contributeurs d'après leurs écrits afin d'améliorer les analyses qui peuvent être réalisées sur les discussions recueillies. Les textes utilisés dans le travail présenté concernent des forums de discussion liés au diabète. Les auteurs présentent une approche utilisant des algorithmes de classification supervisés et s'appuyant sur des lexiques et des ressources terminologiques afin de repérer des caractéristiques telles le genre, la tranche d'âge, la maladie ou le statut du contributeur.

Ce numéro thématique n'aurait pas pu voir le jour sans le travail des relecteurs, que nous remercions ici pour leur regard avisé et leurs recommandations lors de la sélection des articles. Chaque article a été relu par deux relecteurs de la liste ci-dessous et une des éditrices. Nous remercions également les auteurs pour leurs contributions de qualité et pour leur patience, le processus de publication ayant pris plus de temps qu'envisagé.

Jérôme Azé (LIRMM, Montpellier, F)  
 Olivier Bodenreider (NIH, US national Library of Medicine, Bethesda, USA)  
 Sandra Bringay (LIRMM, Montpellier, F)  
 Jean Charlet (SIM, AP-HP, Paris, F)  
 Adrien Coulet (LORIA, Univ. De Lorraine, Nancy, F)  
 Vincent Claveau (IRISA, INRIA, Rennes, F)  
 Olivier Dameron (IRISA, univ. De Rennes 1, Renne, F)  
 Lorraine Goeuriot (IMAG, univ. Grenoble Alpes, Grenoble, F)  
 Natalia Grabar (LIMSI, CNRS, F)  
 Marie-Christine Jaulent (SIM, AP-HP, Paris, F)

NATHALIE AUSSENAC-GILLES  
 IRIT, CNRS, Université de Toulouse  
 UPS – 118, route de Narbonne, 31062 Toulouse Cedex 9  
 Nathalie.aussenac-gilles@irit.fr  
 NATHALIE SOUF  
 IRIT, CUFR J.F. Champollion  
 UPS – 118, route de Narbonne, 31062 Toulouse Cedex 9  
 Nathalie.souf@irit.fr  
 LYNDAMINE-LECHANI  
 IRIT, CNRS, Université Toulouse 3  
 UPS – 118, route de Narbonne, 31062 Toulouse Cedex 9  
 Lynda.tamine-lechani@irit.fr

DOI:10.3166/tsi37-0006 © 2018 Lavoisier