



HAL
open science

Intelligence Artificielle et Santé. Une analyse rétrospective depuis 2010

Jean Charlet, Sandra Bringay

► **To cite this version:**

Jean Charlet, Sandra Bringay. Intelligence Artificielle et Santé. Une analyse rétrospective depuis 2010. IC 2019 - 30es Journées francophones d'Ingénierie des Connaissances, AFIA, Jul 2019, Toulouse, France. pp.26-42. hal-02329573

HAL Id: hal-02329573

<https://hal.science/hal-02329573>

Submitted on 23 Oct 2019

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.

Intelligence Artificielle et Santé. Une analyse rétrospective depuis 2010

Jean Charlet^{1,2}, Sandra Bringay^{3,4}

¹ AP-HP/DRCI, Paris, France

² Sorbonne Université, INSERM, Univ. Paris 13, LIMICS, Paris, France
Jean.Charlet@upmc.fr

³ LIRMM UM CNRS, 34080 Montpellier, France
bringay@lirmm.fr

⁴ AMIS, UNIVERSITÉ PAUL VALÉRY, Montpellier, France

Résumé : Les Journées francophones d'ingénierie des Connaissances ont hébergé de nombreux articles liés à l'intelligence artificielle et à la santé au sein des sessions plénières mais également dans les ateliers associés. Ces journées ont permis de créer des forums d'échanges scientifiques et de créer une communauté de recherche active. L'objectif de cet article est de réaliser une revue semi-automatique des articles produits dans ce contexte sur la thématique de l'intelligence artificielle et de la santé, d'analyser leur impact sur le domaine et d'identifier les défis à relever. Pour cela, nous avons analysé semi-automatiquement 108 articles publiés entre 2010 et 2018, à partir de leurs titre, mots-clés et résumé. Nous avons identifié sept grands thèmes de recherche dont les principaux sont liés à la construction, l'alignement et l'utilisation d'ontologies, la recherche d'informations, les annotations, les recommandations, les bonnes pratiques, le traitement automatique de la langue, la fouille de données de santé et l'analyse des médias sociaux pour des applications liées à la santé. Cette revue montre l'hybridation des approches symboliques et numériques pour la création de méthodes et d'outils au coeur de la médecine du futur.

Mots-clés : santé, intelligence artificielle

1 Introduction

L'intelligence artificielle est une science, dont le but est de faire faire par des machines des tâches que l'homme réalise grâce à son intelligence. Les chercheurs développent pour cela des approches et techniques multiples, en s'appuyant sur diverses disciplines comme l'informatique, les statistiques et la neuroscience mais également la psychologie cognitive, l'ergonomie, la linguistique, la sociologie, en combinant différents types de traitements de la donnée comme le traitement automatique des langues, la construction d'ontologies, la fouille de données, l'apprentissage automatique...

L'intelligence artificielle est au coeur de la médecine du futur. Ses applications concernent en effet toutes les activités humaines, dont la santé. Elles permettent de répondre à de nombreux défis comme l'amélioration de la qualité des soins, avec les opérations assistées, le suivi des patients à distance, les prothèses intelligentes, les traitements personnalisés, grâce au recoupement d'un nombre croissant de données (big data), l'indexation des connaissances, l'aide à la décision, etc.

Pour répondre à ces défis, la communauté de recherche française s'est structurée avec des événements annuels. En particulier, les Journées Francophones d'Ingénierie des Connaissances hébergent chaque année de nombreux articles, liés à la santé au sein des sessions plénières mais également dans des ateliers associés. Ces journées ont permis de créer des forums d'échanges scientifiques et de révéler une communauté de recherche active.

Dans cet article, notre objectif est de réaliser une revue semi-automatique des articles produits sur la thématique de l'intelligence artificielle pour la santé, d'analyser leur impact sur le domaine et d'identifier les défis à relever. Pour cela, nous avons sélectionné *a)* les articles liés à la santé publiés dans la conférence IC depuis 2010 et *b)* les articles des ateliers « IC et

Santé », « IA et Santé » et « Symposium sur l'Ingénierie de l'Information de santé » (SIIM), qui se se sont succédés tous les ans depuis l'année 2011. Pour commencer l'analyse, nous avons utilisé la méthode probabiliste générative non supervisée Latent Dirichlet Allocation (LDA) (Blei *et al.*, 2003). Cette méthode prend en entrée un ensemble de textes et génère automatiquement les thèmes abordés dans ces textes de manière interprétable, sous forme de listes de mots-clés pondérés. LDA a été utilisé pour de nombreuses applications. Dans un contexte similaire, (Wang *et al.*, 2011) et (Wang *et al.*, 2016) ont élaboré un modèle LDA spécialisé utilisant des termes biomédicaux pour explorer la littérature biomédicale. Nous avons identifié six thématiques automatiquement et une dernière manuellement. Ces sept thématiques représentent les centres d'intérêt des chercheurs de la discipline qui hybrident IA symbolique et numérique et ouvrent des perspectives de recherche pertinentes pour améliorer la santé. Notons encore que nous proposons principalement une étude rétrospective et que nous avons fait le choix de ne pas nous projeter dans le futur, principalement pour des raisons de place.

Cet article est organisé de la manière suivante. Dans la section 2, nous décrivons ce qu'est l'IA. Dans la section 3, nous présentons la méthode semi-automatique qui nous a permis de détecter les thèmes ayant émergé dans les publications. Dans la section 4, nous décrivons ces thèmes avant de conclure dans la section 5.

2 Définition de l'IA

L'intelligence artificielle est née dans les années 1950 avec l'objectif de faire produire des tâches humaines par des machines mimant l'activité du cerveau. Face aux déboires des premières heures, deux courants se sont constitués.

Les tenants de l'intelligence artificielle dite forte visent à concevoir une machine capable de raisonner comme l'humain, avec le risque supposé de générer une machine supérieure à l'homme et dotée d'une conscience propre. Cette voie de recherche est toujours explorée aujourd'hui, même si de nombreux chercheurs en IA estiment qu'atteindre un tel objectif est impossible.

D'un autre côté, les tenants de l'intelligence artificielle dite faible mettent en œuvre toutes les technologies disponibles pour concevoir des machines capables d'aider les humains dans leurs tâches. C'est essentiellement cette deuxième approche qui a mobilisé la communauté française travaillant dans le domaine de l'intelligence artificielle et de la santé. Elle a généré de nombreux systèmes spécialisés et performants qui peuplent aujourd'hui notre environnement : créer des profils de futurs « amis » possibles sur les réseaux sociaux, identifier des dates dans les textes pour ordonner des événements dans des comptes-rendus d'hospitalisation, aider le médecin à prendre des décisions, etc. Ces systèmes ont en commun d'être limités dans leurs capacités d'adaptation : ils doivent être manuellement adaptés pour accomplir d'autres tâches que celles pour lesquelles ils ont été initialement conçus. On distingue deux courants, l'IA symbolique et l'IA numérique qui tendent à s'hybrider ces dernières années (Babbar *et al.*, 2018).

2.1 IA symbolique

L'approche la plus ancienne s'appuie sur l'idée que nous raisonnons en appliquant des règles logiques (déduction, classification, hiérarchisation, etc.). Les systèmes conçus sur ce principe appliquent différentes méthodes, fondées sur l'élaboration de modèles d'interaction entre agents (systèmes multi-agents), de modèles syntaxiques et linguistiques (traitement automatique des langues) ou d'élaboration d'ontologies (représentation des connaissances). Ces modèles sont ensuite utilisés par des systèmes de raisonnement logique pour produire des faits nouveaux.

Dans les années 1980, cette approche, dite symbolique, a permis le développement d'outils capables de reproduire les mécanismes cognitifs d'un expert. C'est pourquoi on les a baptisés « systèmes experts ». Les plus célèbres, Mycin (identification d'infections bactériennes) (Buchanan & Shortliffe, 1985) ou Sphinx (détection d'ictères) (Fieschi *et al.*, 1990), s'appuient

sur l'ensemble des connaissances médicales dans un domaine donné et une formalisation des raisonnements des spécialistes qui lient ces connaissances entre elles pour aboutir à un diagnostic.

Les systèmes actuels, qualifiés d'aide à la décision, de gestion des connaissances ou d'e-santé, sont plus sophistiqués. Ils bénéficient de meilleurs modèles de raisonnement ainsi que de meilleures techniques de description des connaissances médicales, des patients et des actes médicaux. La mécanique algorithmique est globalement la même, mais les langages de description sont plus efficaces et les machines plus puissantes. Ils ne cherchent plus à remplacer le médecin, mais à l'épauler dans un raisonnement fondé sur les connaissances médicales de sa spécialité.

2.2 IA numérique

Contrairement à l'approche symbolique, l'approche dite numérique se focalise sur les régularités qu'il est possible d'identifier dans les données pour en extraire des connaissances, sans modèle préétabli. Cette approche, née avec le connexionnisme (Sun, 1999) et les réseaux de neurones artificiels (Rosenblatt, 1957), se développe aujourd'hui grâce à l'augmentation de la puissance des ordinateurs et à l'accumulation des gigantesques quantités de données, le fameux big data.

La plupart des systèmes actuels procèdent par apprentissage automatique, une méthode visant à donner aux ordinateurs la capacité d'apprendre à partir de données des représentations les plus générales ou spécifiques possibles, qui leur permettent de résoudre des tâches sans être explicitement programmés pour chacune. Par exemple, les algorithmes d'apprentissage profond (deep learning) (Bengio, 2009), dont l'usage explose depuis une dizaine d'années, s'inspirent du fonctionnement cérébral. Ils simulent un réseau de neurones organisés en différentes couches, échangeant les uns avec les autres. La force de cette approche est que l'algorithme apprend la tâche qui lui a été assignée par « essais et erreurs », avant de se débrouiller tout seul. On peut citer par exemple les travaux de (Esteva *et al.*, 2017) qui ont proposé un modèle qui sait reconnaître les mélanomes mieux que 20 dermatologues. La raison s'explique par le nombre de cas (d'images) que le modèle a pu apprendre automatiquement très rapidement, ce qui est impossible pour un humain.

2.3 Hybridation entre IA symbolique et numérique

Ce découpage symbolique/numérique est intéressant car il propose une grille d'analyse des différents travaux que nous voulons étudier. Mais ceux-ci, comme nous le verrons dans la section 4, ne se positionnent pas toujours dans une seule case. Des travaux en IA symbolique utilisent les algorithmes de l'IA numérique pour construire les modèles, par exemple, pour construire des ontologies à partir d'algorithmes numériques de fouille de texte. Des travaux en IA numérique surpassent l'état de l'art en intégrant de la connaissance experte issue de méthodes symboliques, par exemple pour des tâches de classification incluant un raisonnement ontologique (Petegrosso *et al.*, 2016; Notaro *et al.*, 2017). Enfin, cette description n'est pas exhaustive et ne tient pas compte d'un champ spécifique de l'IA, en grande partie absente des travaux que nous analysons, la robotique. Ce champ ferait bien partie de notre analyse mais il est assez spécifique et trop peu d'articles à la base de notre étude s'y intéressent.

3 Analyse semi-automatique du corpus d'articles

L'objectif a été de réaliser un premier tri des articles de manière semi-automatique sous la forme de thèmes que nous avons ensuite interprétés et affinés manuellement.

3.1 Préparation des données

Dans ce travail, nous avons composé un jeu de données de 108 articles. Nous avons sélectionné manuellement 34 articles liés à la santé et publiés dans la conférence IC depuis 2010

ainsi que tous les articles des ateliers « IC et Santé », « IA et Santé » et « Symposium sur l'Ingénierie de l'Information de santé » (SIIM), qui se se sont succédés tous les ans depuis l'année 2011. Cinq ont été exclus car rédigés en langue anglaise. Nous avons récupéré pour chaque article le titre, le résumé et les mots-clés qui ont été concaténés dans une chaîne de caractères. Les mots-clés ont été dédoublés pour leur donner plus de poids et certains ont été normalisés manuellement (e.g. alignement et mapping ont été unifiés en alignement, taln et traitement automatique de la langue en TALN, etc.).

3.2 Application du modèle non supervisée LDA pour détecter les thématiques

Plusieurs modèles permettent d'extraire les thèmes d'un corpus textuel comme l'analyse sémantique latente (Landauer & Dutnais, 1997), l'analyse sémantique latente probabiliste (Hofmann, 2001), l'allocation de Dirichlet latente (LDA) (Blei *et al.*, 2003) et l'indexation sémantique latente (Deerwester *et al.*, 1990). Dans cette étude, nous avons utilisé LDA. C'est un modèle probabiliste avec une définition hiérarchique de ses composantes. Il est génératif, ce qui signifie que nous pourrions générer de nouveaux documents à partir d'un modèle donné. Il est basé sur une représentation relativement simple et robuste des documents textuels. Il ne prend pas en compte l'ordre d'occurrence des termes et la structure des phrases. Le principal avantage de LDA est qu'il s'agit d'un modèle probabiliste avec des sujets interprétables. Son principal inconvénient est qu'il n'existe pas de métrique objective qui justifie le choix des hyperparamètres.

Pour un corpus donné de D documents, nous définissons tout d'abord le vocabulaire pertinent V comme une collection prétraitée de termes occurrant dans le corpus. Dans ce travail, les traitements appliqués sont la mise en minuscules, la suppression des mots vides et des ponctuations, la lemmatisation et la génération des bigrammes et trigrammes. Le corpus obtenu est une collection de D documents $D = \{d_1, \dots, d_D\}$. Un document est un N-Uplet de N termes $d = \{t_1, \dots, t_N\}$. À chaque terme $t(d, n)$ est associé un thème représenté par la variable $z(d, n)$. θ_d représente la distribution des thèmes du document d . Des hyperparamètres, α et η , définissent l'a priori sur θ et β où β_k décrit la distribution du thème k .

Pour nos expérimentations, nous avons utilisé les bibliothèques NLTK¹ pour les prétraitements ainsi que Gensim² et pyLDAvis³ pour LDA et sa visualisation.

Le nombre de thèmes en entrée de l'approche est difficile à choisir. Dans ce travail, nous avons fait varier le nombre de thèmes K manuellement et retenu 6 thèmes interprétables.

3.3 Visualisation

La principale information que nous pouvons tirer de l'ajustement d'un modèle LDA sur notre corpus de données textuelles est une structure en thèmes ainsi que la répartition des thèmes dans les documents contenus de ce corpus, et l'association de chaque thème à une liste de mots-clés et leur pondération. La figure 1 correspond à la visualisation obtenue en sortie du modèle LDA. Chaque bulle sur le graphique de gauche représente un thème. Plus la bulle est grande, plus ce thème est important. Un bon modèle aura des bulles assez grandes, ne se chevauchant pas, dispersées dans le graphique au lieu d'être regroupées dans un quadrant. Un modèle comportant trop de thèmes comportera des chevauchements et des bulles de petite taille regroupées dans une région du graphique. En sélectionnant une des bulles, les mots et les barres situés à droite sont mises à jour. Ces mots sont les mots-clés dominants du thème sélectionné. Les barres à côté de ces mots-clés correspondent à leur pondération. La visualisation est disponible en ligne⁴.

1. <https://www.nltk.org/>

2. <https://radimrehurek.com/gensim/>

3. <https://github.com/bmabey/pyLDAvis>

4. <http://www.univ-montp3.fr/miap/sbringay/lda.html>

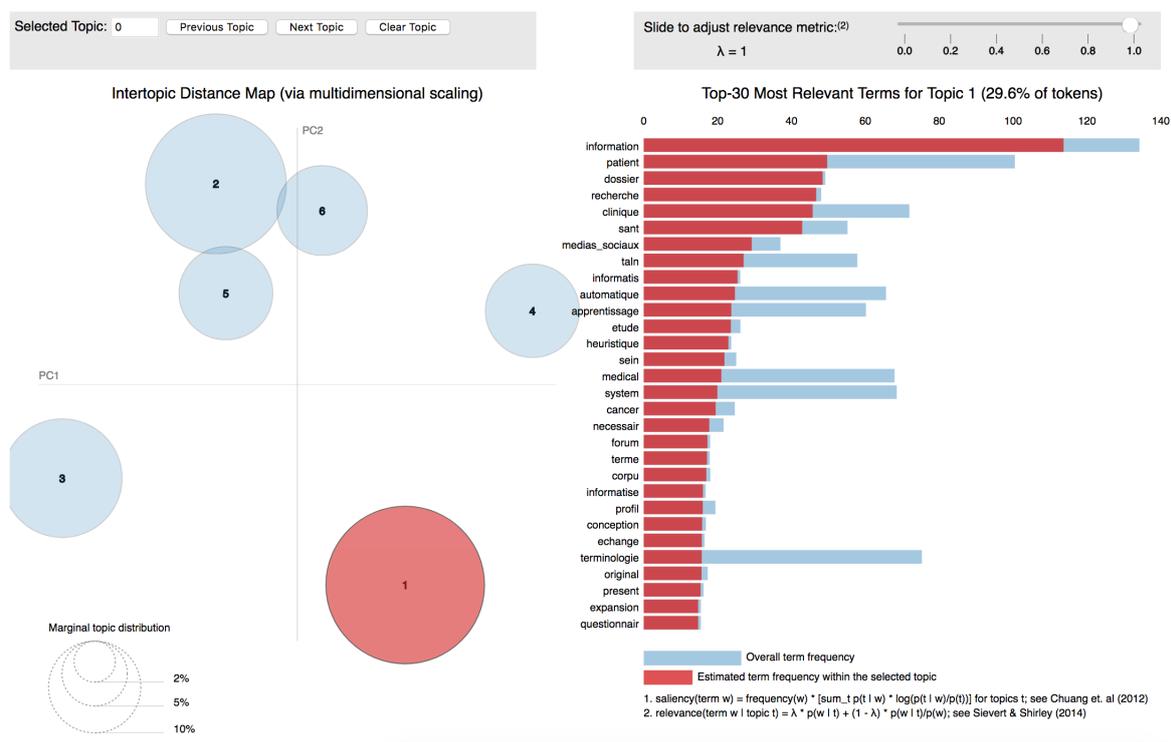


FIGURE 1 – Visualisation des thèmes centrée sur le thème 1 avec les termes associés.

4 Les sept thèmes principaux liés à l'Intelligence Artificielle et la Santé

L'interprétation des thématiques en sortie du modèle LDA a été réalisée manuellement par les deux auteurs de cet article. Parfois, les mots-clés d'un thème ne sont pas suffisants pour donner un sens à un thème. Nous avons donc consulté les documents associés pour en déduire le sens. Nous décrivons dans la suite les 6 thèmes identifiés par la méthode LDA. Un septième thème a été isolé manuellement par les auteurs comme un sous thème du thèmes 2. Par ailleurs, nous avons réintégré manuellement les articles exclus de l'analyse automatique car rédigés en anglais. Pour finir, quatre articles particuliers (deux très généraux et deux très spécifiques) ont donné lieu à un commentaire dans la section 4.8. Il est important de souligner que cette description n'est pas exhaustive et qu'il existe sûrement d'autres thématiques d'intérêt associant IA et Santé en France, donnant lieu à d'autres publications dans d'autres évènements non étudiés dans cet article.

4.1 Thème 1 : construction d'ontologies et de terminologies

La construction d'ontologies pour la santé est depuis longtemps un des domaines d'application de l'Ingénierie des connaissances, avant même la publication des premiers articles pris en compte dans cette analyse. Les articles de ce thème se focalisent sur le processus de construction de ces ontologies avant la description de l'utilisation de l'ontologie même si celle-ci est généralement renseignée dans les deux cas. Une des spécificités françaises, si ce n'est francophone, est, suite aux travaux du groupe TIA (Terminologie et Intelligence Artificielle) d'avoir mis l'accent sur l'utilisation des corpus textuels pour construire les ontologies.

Ce type de démarche a été utilisé pour les travaux suivants, correspondant chacun à l'élaboration d'une ontologie : une ontologie de domaine en médecine périnatale (Dhombres *et al.*, 2010), une ontologie bilingue de la maladie d'Alzheimer (Dramé *et al.*, 2012), une ontologie de la médecine d'urgence pour l'annotation sémantique (Charlet *et al.*, 2012), une ontologie de la psychiatrie centrée sur l'environnement familial et social du patient (Richard *et al.*,

2013) et enfin une ontologie de la sclérose latérale amyotrophique (SLA – Cardoso *et al.* (2017)). Pas très éloignée de cette démarche, on trouve la construction d'une ontologie des infections orthopédiques avec une approche terminologique à partir des termes de UMLS (De Nizza *et al.*, 2013) et la construction d'une terminologie d'interface avec un processus d'élaboration de libellés précis et valides (Nikiema *et al.*, 2017).

D'autres articles vont mettre davantage l'accent sur les connaissances à l'origine de l'ontologie comme la construction d'une ontologie des maladies infectieuses tenant compte du mode de propagation (Camara *et al.*, 2012) ou l'évolution de l'ontologie des urgences sus-nommée en tenant compte du modèle d'une ontologie de l'anatomie, le FMA (*Foundational Model of Anatomy* – Charlet *et al.* (2014)). D'autres encore vont insister sur le caractère modulaire de l'ontologie construite, la plupart du temps, à partir d'une ontologie existante déjà décrite auparavant (Camara *et al.*, 2014; Cardoso *et al.*, 2018).

Dans la même veine que la réflexion sur les connaissances, une autre série d'articles va mettre en avant le modèle des données à traiter pour justifier le développement de l'ontologie. Il en est ainsi d'une ontologie dans le domaine de la résistance des bactéries aux antibiotiques (Assele Kama *et al.*, 2010), d'une première ontologie des maladies rares (Dhombres *et al.*, 2011), d'une ontologie pour la qualification et l'indexation des outils d'analyse en sciences omiques (Henry *et al.*, 2016), ou encore d'une terminologie médicale française pour la détection des médicaments en texte libre (Cossin *et al.*, 2018). Dans ce dernier cas, un modèle de données est développé et une ontologie envisagée.

Enfin, à la limite de l'aide à la décision (Cf. thème 4), un dernier article met en avant la sémiologie quantitative des signes modélisés dans le domaine des traumatismes du genou pour justifier la construction de l'ontologie (Guefack *et al.*, 2012).

4.2 Thème 2 : alignement d'ontologies, interopérabilité

Ce thème met en avant des articles liés aux ontologies. Par rapport au thème 1, plus axé sur la construction de l'ontologie, ce thème met l'accent sur l'utilisation de l'ontologie pour l'alignement et l'interopérabilité, le second n'allant pas sans le premier. En effet, l'interopérabilité nécessite d'avoir travaillé en amont sur les questions d'alignement, de façon manuelle ou (semi-)automatisée.

Dans ce contexte, on trouve des articles sur de l'alignement classique, par rapport aux concepts simples : deux travaux visant le même but, à savoir le développement d'alignements multilingues pour le serveur Bioportal du LIRMM (Annane *et al.*, 2016) ou pour le serveur HeTOP (anciennement CISMEF) du CHU de Rouen (Merabti *et al.*, 2012). Avec des méthodes similaires mais fondées sur d'autres terminologies, (Nikiema *et al.*, 2016) vérifient la possibilité de fournir des alignements fiables alors que (Mary *et al.*, 2016) enrichissent les travaux de la même équipe, 4 ans auparavant (Cf. infra), avec de nombreuses méthodes pour fiabiliser au maximum les alignements proposés. Pour des modèles de données plus complexes, (Maaroufi *et al.*, 2014) proposent de formaliser les alignements et de choisir, en fonction de critères formels, quelle technique utiliser. Nous terminons cette revue de détail sur l'alignement par deux articles d'une même équipe sur un problème qui va devenir de plus en plus important au fur et à mesure de l'utilisation des classifications ou ontologies pour l'annotation sémantique : la maintenance de ces annotations malgré l'évolution des ressources utilisées (Reis *et al.*, 2012; Dinh *et al.*, 2013).

Un groupe d'articles s'intéresse aux stratégies mises en place pour assurer le fonctionnement de serveurs de terminologies HeTOP et Bioportal (Jonquet *et al.*, 2010; Grosjean *et al.*, 2011).

Un autre groupe décrit les modèles d'interopérabilité mis en œuvre, avec des représentations plus ou moins liées au web sémantique, pour croiser des données issues des entrepôts de données (Gaignard *et al.*, 2012; Lelong *et al.*, 2017) ou, dans le cadre des maladies rares, la construction de l'ensemble de données minimum de description des patients pour un serveur permettant de fédérer leurs données (Choquet *et al.*, 2012) ou encore, dans le cadre d'une architecture générale pour le diagnostic médical, la proposition d'une méthode d'intégration d'ontologies (Sow *et al.*, 2018).

Pour terminer, (Traore *et al.*, 2014) abordent la question de l'interopérabilité entre dossiers patients hospitaliers à travers des exemples précis tirés d'un projet ANR. (Raboudi *et al.*, 2017) mettent en avant la question de la traçabilité qui est un concept important pour l'interopérabilité dans la mesure où les données subissent des traitements et des changements qui risquent d'induire des erreurs de transcription et pour lesquels on doit être à même de tracer ces transformations.

4.3 Thème 3 : traitement Automatique de la langue pour le domaine de la santé

Le domaine de la santé est un domaine d'application privilégié pour les méthodes de traitement automatique de la langue naturelle via la création d'outils et de méthodes pour diverses applications visant par exemple l'amélioration des interactions des professionnels de santé avec le dossier patient informatisé via notamment des outils de reconnaissance vocale, l'acquisition de connaissances en matière de santé par les patients, l'identification de patients à partir des textes cliniques, etc.

Un pré-requis à l'analyse des données textuelles des patients est l'anonymisation des identités des personnes et des lieux. (Cardoso *et al.*, 2017) se sont attelés à cette tâche dans le contexte du repérage conceptuel pour la maladie de Charkot. Un autre pré-requis important est la détection de la négation. En effet, les marqueurs de la négation sont constitués d'un ou de plusieurs mots qui modifient la polarité et donc le sens de la phrase (Abdaoui *et al.*, 2017; Dalloux *et al.*, 2017). Pour finir des indices comme la temporalité, le conditionnel ou le possible sont également importants à repérer pour l'interprétation (Garcia-Fernandez *et al.*, 2011).

Une fois les pré-traitements effectués, il est alors possible d'extraire de l'information dans les textes comme des co-occurrences de concepts médicaux (Abdoune *et al.*, 2011), des signes et des symptômes dans le cas des maladies rares (Martin *et al.*, 2014), des événements indésirables médicamenteux (Personeni *et al.*, 2016), etc. Ces différents concepts sont évidemment intéressants à relier entre eux (Minard *et al.*, 2011).

Le dossier patient informatisé est alors une source de données textuelles ayant suscité de nombreux travaux supportant l'annotation (Marrast *et al.*, 2013) mais également des raisonnements à partir de ces annotations (Segond *et al.*, 2014) par exemple pour harmoniser la représentation des comptes rendus et évaluer leur similarité (Parès *et al.*, 2014). D'autres types de données, hors dossiers des patients, ont également été utilisés comme les bulletins pour la santé du végétal (Roussey & Ghorfi, 2018) ou encore les requêtes d'experts médicaux (Znaïdi *et al.*, 2013).

Parmi les récents challenges à relever, on trouve la prise en compte du multi-linguisme dans les documents relatifs à la santé (Cabot *et al.*, 2017) et la génération de contenus (Kamath *et al.*, 2017) pour répondre à des questions médicales.

4.4 Thème 4 : bonnes pratiques et aide à la décision

Ce thème a été isolé manuellement par les auteurs. En effet, l'analyse du thème 2 a rapidement montré qu'il recouvrait beaucoup plus d'articles que les autres. En approfondissant encore l'exploration, on a vu qu'un certain nombre d'articles correspondaient à l'aide à la décision, thématique historique dans la communauté de l'IA appliquée à la médecine et bien connue des auteurs. Cela nous a donc amené à expliciter ce nouveau thème et nous a permis d'y rattacher 11 articles. On pourrait imaginer qu'une nouvelle paramétrisation des algorithmes de LDA nous aurait permis de retrouver ce thème mais nous n'avons plus de temps pour explorer cette hypothèse.

Un premier groupe d'articles montre le développement de systèmes interactifs d'aide à la décision (SIAD) avec des architectures spécifiques : (Chniti *et al.*, 2012) proposent un SIAD dont les concepts manipulés sont des ontologies avec des règles métiers en JRules; (Richard *et al.*, 2018) développent une analyse des besoins des médecins avant de proposer un modèle général de raisonnement médical mettant l'accent sur la sélection des informations à donner aux médecins; (Rybarczyk *et al.*, 2011) développent une plateforme de traitement

multimodal des symptômes aphasiques fondée sur l'analyse des réactions des patients aux divers stimuli.

Un second groupe d'articles propose des SIAD mettant en œuvre des Guides de Bonnes Pratiques (GBP) médicaux : (Meilender *et al.*, 2011) proposent de transformer les recommandations de pratiques cliniques (RPC) en GBP en mettant à disposition un éditeur d'arbre de décision ; (Séroussi *et al.*, 2018) développent un SIAD qui met en œuvre un raisonnement ontologique permettant la réconciliation de différents GBP pour un même patient et présentant les résultats à travers une interface spécifique ; (Schnell *et al.*, 2018) proposent une interprétation des GBP en cancérologie par du raisonnement à partir de cas ; pour finir, (Psiuk & Manet, 2015) analysent la mise en œuvre réelle dans plusieurs hôpitaux du Plan de Soins Type (PST) pour mieux l'implémenter et l'utiliser pour la mise en place de chemins cliniques.

Deux articles analysent des situations de soin pour proposer de nouvelles approches de raisonnement : (Falip *et al.*, 2018) analysent des données médicales de traitement pour proposer un système de recommandation centré sur le patient dans le contexte du raisonnement par analogie ; (Benmouffek *et al.*, 2018) effectuent un travail de recensement et de modélisation de méthodes ayant fait leurs preuves pour l'accompagnement de la prise en charge des personnes avec des troubles neuro-développementaux

Enfin, deux articles abordent des problématiques spécifiques : (Sediri *et al.*, 2012) se fondent sur un modèle de raisonnement pour la gestion de crise pour définir une structure de retour d'expériences et les interfaces permettant de gérer la crise ; (Ugon, 2018) mettent en évidence la difficulté d'évaluer les algorithmes de scoring automatique des données médicales temporelles par rapport aux choix des indicateurs et à la définition des bons *gold standard* auquel se comparer.

4.5 Thème 5 : recherche d'informations, annotations et recommandations

Si le domaine de la recherche d'informations tel que nous le concevons aujourd'hui a émergé dans les années 1950, celui-ci a été rapidement spécialisé dans le domaine de la santé du fait de la nature particulière des documents à indexer et du vocabulaire employé dans ces documents.

Le développement de terminologies et d'ontologies dans le domaine de la santé a permis la mise en place de nouvelles méthodes d'indexation sémantique. La question des modèles a tout d'abord été centrale. Par exemple, (Dinh & Tamine, 2011) ont développé un modèle de recherche d'information multi-terminologique dédié aux documents biomédicaux. (Ghezaiel *et al.*, 2011) ont développé un réseau proxémique pour la recherche d'informations spécifiques à la maladie des dystonies. (Ranwez *et al.*, 2015) ont proposé l'ontologie OntoToxnucl pour représenter la toxicologie nucléaire environnementale.

L'automatisation des annotations sémantiques des textes à partir de ces modèles, est une condition au développement de systèmes de recherche d'information pertinents. Ces annotations sémantiques viennent enrichir les textes et peuvent donc être exploitées ultérieurement comme un point d'accès à la sémantique de ces textes et interrogées sous la forme de divers traitements. (Brut *et al.*, 2011) se sont ainsi intéressés à l'annotation sémantique du dossier patient informatisé. (Dayre & Batatia, 2011) se sont focalisés sur l'annotation de médias selon un modèle de l'activité. (Azzi *et al.*, 2016) ont exploité des annotations sémantiques pour automatiser le calcul des valeurs nutritionnelles d'une recette de cuisine.

Si les textes sont généralement l'objet de la recherche d'information, d'autres médias peuvent représenter la cible de la recherche comme les images dans le système d'information par modalités développé par (Tirilly, 2011) ou encore la recherche sémantique d'images en gastroentérologie (Chabane & Rey, 2013). Une difficulté supplémentaire peut également venir du croisement de différents types de données par exemple cliniques et omiques dans le dossier patient informatisé (Cabot *et al.*, 2015). Pour finir, la disponibilité des données dans le cas de réseaux dégradés (Azanzi *et al.*, 2012) peut également poser la question de l'intégration de différentes sources de données interopérables.

Dernièrement, la recherche d'information a tiré parti des méthodes d'apprentissage pour capturer la sémantique des documents (Nguyen *et al.*, 2017) et raisonner sur ces documents, par exemple pour le codage automatique du PMSI (Ternois *et al.*, 2018), la génération d'alertes

lors de la prescription (Lindemann, 2018), la description de cohortes guidée par les données (Neuraz *et al.*, 2018) ou encore le *rescreening* à partir d'un entrepôt de données cliniques (Pasco *et al.*, 2018).

4.6 Thème 6 : fouille des données de santé

Les méthodes classiques de fouille de données ont été utilisées pour extraire de la connaissance nouvelle à partir de grandes quantités de données de santé. Cette connaissance peut alors être utilisée pour enrichir les interprétations des professionnels de la santé, tout en fournissant des entrées pour des méthodes automatiques ou semi-automatiques exploitant cette connaissance.

On trouve tout d'abord les approches de type recherche de motifs qui vont repérer des régularités généralement fréquentes dans les données : par exemple, (Béchet *et al.*, 2012) utilisent des motifs séquentiels pour découvrir des relations entre des gènes et des maladies rares dans la littérature biomédicale; (Pinaire *et al.*, 2015b) utilisent ces motifs pour représenter des parcours de soins à partir de données du PMSI; (Personeni *et al.*, 2018) découvrent des associations entre évènements indésirables médicamenteux.

Si ces motifs sont intéressants pour représenter des données ordonnées (des évènements dans le temps ou des mots dans des données séquentielles textuelles), ils peuvent être étendus pour intégrer de nouvelles dimensions comme, par exemple, la dimension spatiale dans le cas des trajectoires (Pinaire *et al.*, 2017b). Ces motifs ont été utilisés afin de prédire l'évolution des populations de patients et les coûts associés et ainsi recommander les soins les plus efficaces et les moins coûteux (Pinaire *et al.*, 2017a).

D'autres auteurs ont utilisé l'analyse formelle de concepts pour étudier des concepts médicaux lorsqu'ils sont décrits formellement, par exemple pour construire des profils de patients (Séroussi *et al.*, 2013).

D'autres types de données ont également été utilisés comme des images (Abbal *et al.*, 2011) en imagerie ultrasonore ou encore des données textuelles pour parcourir la littérature afin de produire un état de l'art semi-automatique (Thebault *et al.*, 2010; Pinaire *et al.*, 2016).

Les régularités extraites à partir des gros volumes de données médicales vont alors pouvoir être utilisées pour des tâches de prédiction par exemple pour l'aide au diagnostic de tumeurs cérébrales (Ben Lamine *et al.*, 2012), la prédiction des stades de sommeil (Ugon *et al.*, 2016), la réadmission à l'hôpital (Bussy *et al.*, 2018), le diagnostic du diabète (Dendani & Allouani, 2018) ou encore la biométrie de la tête du fœtus (Grandjean *et al.*, 2018).

La visualisation des résultats des méthodes de fouille de données est particulièrement importante pour une bonne appropriation des outils et une interprétabilité des approches par les professionnels de santé. Par exemple, (Pinaire *et al.*, 2015a) ont proposé une visualisation des trajectoires de soins des patients, (Lamy *et al.*, 2017, 2015) une plateforme pour comparer les caractéristiques de médicaments, (Serres *et al.*, 2011) un outil pour le suivi des dissections ou encore pour la visualisation de données multi-sources de chimiothérapie (Jannot *et al.*, 2018).

4.7 Thème 7 : Analyse des médias sociaux

L'extraction de connaissances à partir des données hétérogènes issues des médias sociaux de santé pour des applications liées à la santé est un des thèmes ayant animé la communauté ces dernières années. Ces travaux permettent de prendre en compte de manière originale, les données produites directement par les patients par opposition à des approches où seules les données médicales produites par les professionnels de santé sont étudiées.

Un cadre fédérateur de représentation de ces données, de leur contexte et des indices que l'on peut retrouver dans les médias sociaux a été proposé par (Bricon-Souf *et al.*, 2015; Chahbandarian *et al.*, 2014).

D'autres approches s'intéressent aux thématiques d'intérêt des patients avec des approches d'apprentissage non supervisées sans *a priori* (Pertin *et al.*, 2017) et des approches supervisées nécessitant au préalable la définition de classe d'intérêts (Opitz *et al.*, 2014). L'analyse de ces contenus permet d'identifier des textes liés au risque (Mercadier *et al.*, 2018), à l'incertitude (Abdaoui *et al.*, 2014), aux sentiments (Bringay *et al.*, 2014), à la détresse psycho-

logique (Kessler *et al.*, 2018) ou aux comportements suicidaires (Combes *et al.*, 2016), etc. Ces informations permettent aux professionnels de santé de mieux cerner les perceptions que les patients ont de leur maladie.

Un point récurrent dans ces travaux est la difficulté à mettre en œuvre les chaînes de traitements classiques sur les textes rédigés par les patients qui sont par nature hétérogènes, incertains, entachés d'erreurs et qui nécessitent par conséquent de nombreux pré-traitements. (Tapi Nzali *et al.*, 2015) ont pour cela construit une nouvelle ressource mettant en relation le vocabulaire des patients et celui des professionnels de santé. Cette ressource a été intégrée dans le portail Bioportal⁵ qui permet d'annoter des textes de patients avec des concepts médicaux (Eholié *et al.*, 2016).

Pour finir, les aspects temporels ont également été étudiés pour capturer l'évolution de comportements ou de thématiques au cours du temps selon l'histoire des patients. Ainsi, (Moulahi *et al.*, 2017) ont pu mettre en avant des sujets d'interrogation des patients associés à des états d'esprit et à un événement de leur histoire médicale, jusqu'alors méconnus des professionnels de santé dans le cas du suivi et de la détection des idées suicidaires.

4.8 Articles particuliers

Reste en dehors des thèmes identifiés (6+1), quatre articles, de caractères plus généraux ou avec des spécificités précises. Dès l'instant qu'il apparaissait des articles difficiles à classer, nous n'avons pas cherché à faire disparaître ce « 8^e » thème. Dedans, se trouvent deux articles généraux, celui sur la définition de l'IA par un des auteurs de cet article (Charlet, 2018) et un article qui interroge la gestion des connaissances médicales par rapport à des expériences de terrain (Blanchet *et al.*, 2015). Les deux derniers articles abordent des thématiques très spécifiques : le premier s'intéresse à l'efficacité de la traduction pour améliorer l'accès à l'information médicale dans un contexte transfrontalier (Laforest *et al.*, 2011) et le second teste l'efficacité de routines de déidentification et d'exportation d'images (Seymour & Payoux, 2017).

5 Conclusion, Discussions et Perspectives

Dans cet article, nous avons présenté une étude préliminaire sur les interactions entre le domaine de l'intelligence artificielle et celui de la santé en France. Pour réaliser ce travail, nous avons utilisé une méthode classique de science des données, LDA, pour extraire les principaux thèmes d'intérêts des publications depuis 2010 dans la conférence Ingénierie des connaissances et dans les ateliers associés à cette conférence. Nous avons identifié sept principaux thèmes qui montrent la variété des travaux réalisés ainsi que l'hybridation entre les approches symboliques et numériques. Nous avons aussi fait le choix de ranger chaque article dans un thème unique. Certains d'entre eux pourraient être associés à plusieurs thèmes, en plus ou à la place de celui choisi mais notre but était de mettre en avant les thématiques en prenant chaque article comme exemple illustratif d'une thématique unique.

La principale limitation de cette étude est que nous avons utilisé LDA uniquement comme une entrée à l'interprétation manuelle. LDA nécessite beaucoup de réglages manuels des paramètres. Nous avons passé beaucoup de temps à les identifier pour que les résultats puissent être interprétés de manière significative. En effet, il n'existe pas de métrique objective qui justifie le choix de ces paramètres et en particulier le nombre de thèmes K utilisés en entrée de la méthode, ce qui rend très difficile la généralisation de l'approche à d'autres données et tâches. Par exemple, le thème sur l'aide à la décision a été créé par les auteurs pour reclasser un certain nombre d'articles *a posteriori* du découpage en thèmes fourni par LDA. Par ailleurs, les articles écrits en anglais ont été reclassés à la main. Il est évident également que certains thèmes se recoupent, par exemple le TAL est utile pour la construction d'ontologies ou encore leur alignement. Enfin, cette étude se limite à une seule source d'articles alors que d'autres articles en français liés à l'Intelligence Artificielle et la santé seraient pertinents à

5. <http://bioportal.lirmm.fr/>

analyser. Une autre perspective consiste à reproduire l'analyse en travaillant sur des résumés en langue anglaise obtenus via l'API pubmed⁶ à partir des mots-clés traduits repérés dans les articles français afin de voir si les tendances se confirment au niveau international. Dans ce contexte, le *YearBook of Medical Informatics* qui paraît tous les ans serait une excellente source de textes à analyser puisqu'il est découpé en chapitres précis correspondant pour la plupart à des thématiques de l'IA, comme par exemple (Dhombres & Charlet, 2018). Pour finir, il semble également important de prendre en compte des aspects temporels pour établir une chronologie de l'évolution des thèmes.

Références

- ABBAL R., BASARAB A. & KOUAMÉ D. (2011). Estimation de décalages subpixeliques en imagerie ultrasonore. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 111–118, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-10.pdf>.
- ABDAOUI A., AZÉ J., BRINGAY S., PONCELET P. & GRABAR N. (2014). Analyse des messages des patients et des médecins dans les fora de santé. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Abdaoui.pdf>.
- ABDAOUI A., TCHECHMEDJIEV A., DIGAN W., BRINGAY S. & JONQUET C. (2017). French ConText : Détecter la négation, la temporalité et le sujet dans les textes cliniques Français. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 7–16, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper7.pdf.
- ABDOUNE H., SOUALMIA L. & JOUBERT M. (2011). Analyse de cooccurrences de concepts biomédicaux dans. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 71–76, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-6.pdf>.
- ANNANE A., ÉMONET V., AZOUAOU F. & JONQUET C. (2016). Réconciliation d'alignements multilingues dans BioPortal. In N. PERNELLE, Ed., *27èmes Journées Francophones d'Ingénierie des Connaissances - IC 2016*, Montpellier, France.
- ASSELE KAMA A., MELS G., CHOQUET R., CHARLET J. & JAULENT M.-C. (2010). Une approche ontologique pour l'exploitation de données cliniques. In S. DESPRÈS, Ed., *21èmes Journées Francophones d'Ingénierie des Connaissances - IC 2010*, p. 183–194, Nîmes, France : Ecole des Mines d'Alès.
- AZANZI F. J., LO M. & TCHUENTE M. (2012). Vers une approche d'intégration de données adaptée aux réseaux dégradés : application au système d'information sanitaire camerounais. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- AZZI R., SYLVIE D. & NOBECOURT J. (2016). Approche sémantique pour automatiser le calcul des valeurs nutritionnelles d'une recette de cuisine. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- BABBAR P., SINGHAL A., YADAV K. & SHARMA V. (2018). Connectionist model in artificial intelligence. *International Journal of Applied Engineering Research*, **13**(7), 5154–5159.
- BÉCHET N., CELLIER P., CHARNOIS T. & CRÉMILLEUX B. (2012). Fouille de motifs séquentiels pour la découverte de relations entre gènes et maladies rares. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, p. 149–164, Paris, France.
- BEN LAMINE F. F., KALTI K. & MAHJOUB M. A. M. (2012). Etude de Modèles à base de réseaux Bayésiens pour l'aide au diagnostic de tumeurs cérébrales. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, Paris, France.
- BENGIO Y. (2009). Learning deep architectures for ai. *Found. Trends Mach. Learn.*, **2**(1), 1–127.
- BENMOUFFEK D., PACINI L., HONION J., REYDON H. & KERBIRIOU C. (2018). Données de prise en charge pluridisciplinaire des personnes avec TND. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 75–80, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- BLANCHET P., REUN R. L. & MORVAN T. (2015). Partage de connaissances médicales par les systèmes automatisés : les besoins des acteurs de terrain. In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 32–42, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- BLEI D. M., NG A. Y. & JORDAN M. I. (2003). Latent dirichlet allocation. *J. Mach. Learn. Res.*, **3**, 993–1022.
- BOUAUD J., SAUQUET D., GIRAL P., JULIEN J., CORNET P., FALCOFF H. & SÉROUSSI B. (2010). Pourquoi les médecins ne suivent-ils pas les systèmes de recommandations de bonnes pratiques ? Une hypothèse liée à l'utilisabilité évaluée avec le mode guidé d'ASTI.
- BOUAUD J. & SÉROUSSI B. (2012). Médecine factuelle et recommandations de bonne pratique : une extension du modèle classique pour expliquer les décisions médicales non conformes. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, p. 251–266, Paris, France.
- BOUAUD J., SOULET A., SPANO J.-P., LEFRANC J.-P., COJEAN-ZELEK I., BLASZKA-JAULERRY B., ZE-LEK L., DURIEUX A., TOURNIGAND C., MESSAI N., ROUSSEAU A. & SÉROUSSI B. (2014). Quels sont les patients atteints d'un cancer du sein dont la décision de prise en charge thérapeutique bénéficie de

6. <https://www.ncbi.nlm.nih.gov/pubmed/>

- l'utilisation d'un système d'aide à la décision? Un exemple utilisant la fouille de données et OncoDoc2. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 107–118, Clermont-Ferrand, France. Session 2 : Utilisateurs et usages.
- BRICON-SOUF N., CHANBANDARIAN G. & HO-DAC M. (2015). Un cadre fédérateur de représentation des données et indices issus de forums de santé. In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 43–48, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- BRINGAY S., KERGOSIEN E., POMPIDOR P. & PONCELET P. (2014). Identifier la cible des émotions dans les forums de santé. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 163–174, Clermont-Ferrand, France. Session 3 : Web social.
- BRUT M., AL-KUKHUN D., PÉNINO A., CANUT M.-F. & SÈDES F. (2011). Structuration et Accès au Dossier Médical Personnel : approche par ontologies et politiques d'accès XACML. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 77–86, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-7.pdf>.
- B. G. BUCHANAN & E. H. SHORTLIFFE, Eds. (1985). *Rule-Based Expert Systems : The MYCIN Experiments of the Stanford Heuristic Programming Project*. Reading, MA : Addison-Wesley.
- BUSSY S., VEIL R., LOOTEN V., BURGUN A., GAÏFFAS S., GUILLOUX A., RANQUE B. & JANNOT A.-S. (2018). Design d'un algorithme d'IA en grande dimension pour prédire la réadmission à l'hôpital. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 1–7, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- CABOT C., SOUALMIA L. F. & DARMONI S. J. (2015). Intégration de données cliniques et omiques pour la recherche d'information dans le Dossier Patient Informatisé. In M.-H. ABEL, Ed., *26èmes Journées Francophones d'Ingénierie des Connaissances - IC 2015*, Rennes, France.
- CABOT C., SOUALMIA L. F. & DARMONI S. J. (2017). CIM-IND : Un système multilingue pour l'extraction d'information dans les textes cliniques. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 105–112, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper1.pdf.
- CAMARA G., DESPRÉS S., DJEDIDI R. & LO M. (2012). Vers une ontologie des processus de propagation des maladies infectieuses. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, p. 99–114, Paris, France.
- CAMARA G., DESPRÉS S. & LO M. (2014). IDOSCHISTO : une extension de l'ontologie noyau des maladies infectieuses (IDO-Core) pour la schistosomiase. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 39–50, Clermont-Ferrand, France. Session 1 : Construction, peuplement et exploitation d'ontologies.
- CARDOSO S., AIME X., MEININGER V., GRABLI D., COHEN K. B. & CHARLET J. (2018). De l'intérêt des ontologies modulaires. Application à la modélisation de la prise en charge de la SLA. In S. RANWEZ, Ed., *29èmes Journées Francophones d'Ingénierie des Connaissances - IC 2018*, p. 121–128, Nancy, France : AFIA.
- CARDOSO S., AIME X., MORA L. F. M., GRABLI D., MEININGER V. & CHARLET J. (2016). Les ontologies pour aider à comprendre les parcours de santé dans le cadre des maladies neurodégénératives. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- CARDOSO S., MELO MORA L. F., JAULENT M.-C., AIME X., GRABLI D., MEININGER V. & CHARLET J. (2017). Ontologie et TALN : l'anonymisation au service du repérage conceptuel dans le contexte de la SLA. In C. ROUSSEY, Ed., *28èmes Journées Francophones d'Ingénierie des Connaissances - IC 2017*, p. 98–103, Caen, France.
- CHABANE Y. & REY C. (2013). Annotation et recherche sémantique d'images en gastroentérologie. In L. TAMINE-LECHANI & L. SOUALMIA, Eds., *Actes de SIIM 2013, 2e Symposium sur l'Ingénierie de l'Information Médicale*, Lille. https://www.irit.fr/SIIM/2013/03_siim13.pdf.
- CHAHBANDARIAN G., BRICON-SOUF N., BASTIDE R. & STEINBACH J.-C. (2017). Stable Feature Selection Approach : Application to the Encoding of Secondary Diagnoses. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 77–84, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper11.pdf.
- CHAHBANDARIAN G., MOJAHID M. & BRICON-SOUF N. (2014). Contextual presentation of medical forum's discussions. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Chahbandarian.pdf>.
- CHARLET J. (2018). IA en santé. Définitions, réalisations et perspectives. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 9–15, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- CHARLET J., DECLERCK G., DHOMBRES F., GAYET P., MIROUX P. & VANDENBUSSCHE P.-Y. (2012). Construire une ontologie médicale pour la recherche d'information : problématiques terminologiques et de modélisation. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, p. 33–48, Paris, France.
- CHARLET J., MAZUEL L., DECLERCK G., MIROUX P. & GAYET P. (2014). Décrire les maladies localisées et la physiopathologie pour une ontologie des urgences : un algorithme générique à partir de la FMA. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 15–26, Clermont-Ferrand, France. Session 1 : Construction, peuplement et exploitation d'ontologies.

- CHNITI A., BOUSSADI A., ALBERT P. & CHARLET J. (2012). Validation pharmaceutique des prescriptions médicamenteuses à base d'une ontologie OWL et de règles métier. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- CHOQUET R., MESSIAEN C., PRIOUZEAU A. & DE CARRARA A. (2012). Un jeu de données minimum pour faciliter l'interopérabilité des bases de données pour les maladies rares. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- COMBES P., COMBES S. & MONZIOLS M. (2016). Tentatives de suicide, prédire la récurrence avec des techniques d'apprentissage statistique. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- COSSIN S., LOUSTAU R., JOUHET V., LÉTINIER L., MOUGIN F., EVRARD G., GIL-JARDINÉ C., DIALLO G. & THIESSARD F. (2018). ROMEDI, une terminologie médicale française pour la détection des médicaments en texte libre. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 17–22, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- DALLOUX C., GRABAR N. & CLAVEAU V. (2017). Détection de la négation : corpus français et apprentissage supervisé. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 17–24, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_pape17.pdf.
- DAYRE P. & BATATIA H. (2011). Annotation collaborative de médias pour l'émergence et l'apprentissage de concepts dans le milieu médical. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 133–140, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-12.pdf>.
- DE NIZZA D., ORTIZ J., MEURISSE H. & SCHOBENS P.-Y. (2013). Formalisation et Construction d'une Ontologie dans le Domaine des Infections Orthopédiques. In R. TRONCY, Ed., *24èmes Journées Francophones d'Ingénierie des Connaissances - IC 2013*, Lille, France.
- DEERWESTER S., DUMAIS S. T., FURNAS G. W., LANDAUER T. K. & HARSHMAN R. (1990). Indexing by latent semantic analysis. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE*, 41(6), 391–407.
- DENDANI N. & ALLOUANI R. (2018). A decision support system for the diagnosis of the Diabetes disease. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 81–85, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- DHOMBRES F. & CHARLET J. (2018). As ontologies reach maturity, Artificial Intelligence starts being fully efficient : Findings from the section on knowledge representation and management for the yearbook 2018. 27(1), 140–145. <https://www.ncbi.nlm.nih.gov/pubmed/30157517>.
- DHOMBRES F., JOUANNIC J.-M., JAULENT M.-C. & CHARLET J. (2010). Choix méthodologiques pour la construction d'une ontologie de domaine en médecine périnatale. In S. DESPRÈS, Ed., *21èmes Journées Francophones d'Ingénierie des Connaissances - IC 2010*, p. 183–194, Nîmes, France : Ecole des Mines d'Alès.
- DHOMBRES F., VANDENBUSSCHE P.-Y., RATH A., HANAEUR M., OLYR A., URBERO B., CHOQUET R. & CHARLET J. (2011). Projet OrphaOnto - Première étape de l'ontologisation des bases de connaissances d'Orphanet. In A. MILLE, Ed., *22èmes Journées Francophones d'Ingénierie des Connaissances - IC 2011*, p. 573–588, Chambéry, France.
- DINH D., REIS J. C. D., SILVEIRA M. D. & PRUSKI C. (2013). Identification des informations conceptuelles définissant un alignement entre ontologies médicales. In L. TAMINE-LECHANI & L. SOUALMIA, Eds., *Actes de SIIM 2013, 2e Symposium sur l'Ingénierie de l'Information Médicale*, Lille. https://www.irit.fr/SIIM/2013/05_siim13.pdf.
- DINH D. & TAMINE L. (2011). Vers un modèle de recherche d'information multi-terminologique des documents biomédicaux. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 43–56, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-4.pdf>.
- DRAMÉ K., DIALLO G. & MOUGIN F. (2012). Construction d'une ontologie bilingue de la maladie d'Alzheimer à partir de textes médicaux. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- EHOLIÉ S., NZALI M. D. T., BRINGAY S. & JONQUET C. (2016). MuEVo, un vocabulaire multi-expertise (patient/médecin) dédié au cancer du sein. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- ESTEVA A., KUPREL B., NOVOA R. A., KO J., SWETTER S. M., BLAU H. M. & THRUN S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115.
- FALIP J., BLANCHARD F. & HERBIN M. (2018). Exploration et système de recommandation pour l'aide au raisonnement médical. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 87–91, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- FIESCHI M., DING Y., TANG J., DONG X., HE B. & QIU J. (1990). The sphinx system environment. *Artificial Intelligence in Medicine*.
- FRANDJI B. & JAULENT M.-C. (2012). Un SADC intégré et interopérable dans le Système d'Information Clinique. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- GAIGNARD A., MONTAGNAT J., ZUCKER C. F. & CORBY O. (2012). Fédération multi-sources en neu-

- rosiences : intégration de données relationnelles et sémantiques. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- GARCIA-FERNANDEZ A., LIGOZAT A.-L. & BERNHARD D. (2011). Présent, hypothétique, conditionnel ? Annotation du statut des problèmes médicaux dans des comptes-rendus cliniques en français. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 3–14, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-1.pdf>.
- GHEZAIEL L. B., LATIRI C., AHMED M. B. & GOUIDER-KHOUBA N. (2011). Un réseau proxémique pour la recherche d'information : Application à la maladie des dystonies. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 25–42, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-3.pdf>.
- GRANDJEAN G. A., HOSSU G., BERTHOLDT C., NOBLE P., MOREL O. & GRANGÉ G. (2018). Artificial intelligence assistance for fetal head biometry : Assessment of automated measurement software. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 709–716, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- GROSJEAN J., SOUALMIA L. F., MERABTI T., DAHAMNA B., KERGOURLAY I., THIRION B. & DARMONI S. J. (2011). The French health multi-terminology portal. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 57–70, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-5.pdf>.
- GUEFACK V. D., BERTAUD-GOUNOT V., BURGUN A., DUVAUFERRIER R. & LASBLEIZ J. (2012). Ontologie sémiologique biomédicale et sémiologie quantitative. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- HENRY V. J., SOUALMIA L. F., GROSJEAN J., DESFEUX A., DARMONI S. J. & GONZALEZ B. J. (2016). Omiconto : une ressource termino-ontologique pour la qualification et l'indexation des outils d'analyse en sciences omiques. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- HOFMANN T. (2001). Unsupervised learning by probabilistic latent semantic analysis. *Mach. Learn.*, **42**(1-2), 177–196.
- JANNOT A.-S., ZAPLETAL, BARBIERI A., GEROLDINGER A., BOULET S., ZOHAR S. & ZANAN A. (2018). Intégration et synthèse visuelle de données multi-sources et hétérogènes de chimiothérapie. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 23–29, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- JONQUET C., COULET A., SHAH N. & MUSEN M. (2010). Indexation et intégration de ressources textuelles à l'aide d'ontologies : application au domaine biomédical. In S. DESPRÈS, Ed., *21èmes Journées Francophones d'Ingénierie des Connaissances - IC 2010*, p. 183–194, Nîmes, France : Ecole des Mines d'Alès.
- KAMATH S., GRAU B. & MA Y. (2017). A Study of Word Embeddings for Biomedical Question Answering. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 35–38, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper12.pdf.
- KESSLER R., BÉCHET N., LEDEGEN G. & PUGNIERE-SAAVEDRA F. (2018). Exploration par apprentissage de discussions de personnes en détresse psychologique. In S. RANWEZ, Ed., *29es Journées Francophones d'Ingénierie des Connaissances, IC 2018*, p. 95–102, Nancy, France.
- LAFOREST F., FLORY A. & GARIN-MICHAUD A. (2011). Accès transfrontalier aux informations médicales : un système de traduction pour le projet européen ALIAS. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 5–23, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-2.pdf>.
- LAMY J.-B., SOUALMIA L. F., VENOT A. & DUCLOS C. (2014). Validation de la sémantique d'un langage iconique médical à l'aide d'une ontologie : méthodes et applications. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 51–62, Clermont-Ferrand, France. Session 1 : Construction, peuplement et exploitation d'ontologies.
- LAMY J.-B., UGON A., DUCLOS C., VENOT A., FAVRE M. & BERTHELOT H. (2017). Une plate-forme visuelle pour une information comparative sur les nouveaux médicaments. In C. ROUSSEY, Ed., *28èmes Journées Francophones d'Ingénierie des Connaissances - IC 2017*, p. 38–49, Caen, France.
- LAMY J.-B., UGON A., FAVRE M., VENOT A. & BERTHELOT H. (2015). Comparaison et visualisation des contre-indications des médicaments. In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 26–33, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- LANDAUER T. K. & DUTNAIS S. T. (1997). A solution to plato's problem : The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *PSYCHOLOGICAL REVIEW*, **104**(2), 211–240.
- LELONG R., SOUALMIA L., SAKJI S., DAHAMNA B. & DARMONI S. (2017). Une technologie NoSQL au service de moteur de recherche en santé. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 41–48, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper2.pdf.
- LINDEMANN W. B. (2018). On the quality of automatic alerts during electronic prescription and the possibilities of improvement. In J. CHARLET, M.-D. DEVIGNES & B. SÉROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 93–99, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- MAAROUI M., CHOQUET R. & LANDAIS P. (2014). Formalisation des correspondances pour l'optimisation des alignements automatisés de schémas de données : Application au domaine des maladies rares.

- In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Maaroufi.pdf>.
- MARRAST P., SOUF N. & ZARATÉ P. (2013). Heuristiques et annotations pour (re)penser le Dossier Clinique Informatisé. In L. TAMINE-LECHANI & L. SOUALMIA, Eds., *Actes de SIIM 2013, 2e Symposium sur l'Ingénierie de l'Information Médicale*, Lille. https://www.irit.fr/SIIM/2013/01_siim13.pdf.
- MARTIN L., BATTISTELLI D. & CHARNOIS T. (2014). Mise en place d'une méthode de reconnaissance des signes et des symptômes dans le contexte des maladies rares. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Martin.pdf>.
- MARY M., SOUALMIA L. F. & GANSEL X. (2016). Interopérabilité sémantique dans le domaine du diagnostic in vitro. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- MEILENDER T., JAY N., LIEBER J. & PALOMARES F. (2011). Édition sémantique d'arbres de décision pour l'oncologie avec KCATOS. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 121–132, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-11.pdf>.
- MERABTI T., SOUALMIA L. F., GROSJEAN J., JOUBERT M. & DARMONI S. J. (2012). Méthodes d'alignement de terminologies médicales et leur intégration dans un portail. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icasante2012/programme>.
- MERCADIER Y., AZÉ J., BRINGAY S., CLAVIER V., CUENCA E., PAGANELLI C., PONCELET P. & SALLABERRY A. (2018). #AIDS Analyse Information Dangers Sexualité : caractériser les discours à propos du VIH dans les forums de santé. In S. RANWEZ, Ed., *29èmes Journées Francophones d'Ingénierie des Connaissances - IC 2018*, p. 71–86, Nancy, France : AFIA.
- MINARD A.-L., LIGOZAT A.-L. & GRAU B. (2011). Extraction de relations dans des comptes rendus hospitaliers. In A. MILLE, Ed., *22èmes Journées Francophones d'Ingénierie des Connaissances - IC 2011*, p. 491–506, Chambéry, France.
- MOULAH B., AZÉ J. & BRINGAY S. (2017). Suivi et détection des idéations suicidaires dans les médias sociaux. In C. ROUSSEY, Ed., *28èmes Journées Francophones d'Ingénierie des Connaissances - IC 2017*, p. 26–37, Caen, France.
- NEURAZ A., GARCELON N., BURGUN A. & RANCE B. (2018). multiWAS : interactive and multimodal phenome-wide scan for data-driven cohort description. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 31–36, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- NGUYEN G.-H., TAMINE L., SOULIER L. & SOUF N. (2017). Apprentissage de représentation des documents médicaux guidé par les concepts pour la recherche d'information. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 27–34, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_papers.pdf.
- NIKIEMA J. N., JOUHET V. & MOUGIN F. (2016). Evaluation de la SNOMED CT comme support à l'alignement de terminologies diagnostiques en cancérologie. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- NIKIEMA J. N., MOUGIN F. & JOUHET V. (2017). Processus de prétraitement des libellés d'une terminologie d'interface. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p.9, Toulouse.
- NOTARO M., SCHUBACH M., ROBINSON P. N. & VALENTINI G. (2017). Prediction of Human Phenotype Ontology terms by means of hierarchical ensemble methods. *BMC Bioinformatics*, **18**(1).
- OPITZ T., AZÉ J., BRINGAY S., JOUTARD C. & MOLLEVI C. (2014). Paroles de patients dans les forums de santé : une perspective originale sur la qualité de la vie. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Opitz.pdf>.
- PARÈS Y., AIMÉ X., CHARLET J. & JAULENT M.-C. (2014). Vers une harmonisation automatique de la représentation de comptes rendus médicaux pour évaluer leurs similarités. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Pares.pdf>.
- PASCO J., CAMPILLO-GIMENEZ B., GRAMMATICO-GUILLON L. & CUGGIA M. (2018). Pré-screening en cancérologie : automatisation à partir des entrepôts de données cliniques. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 37–42, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- PERSONENI G., BRESSO E., DEVIGNES M.-D., SMAIL-TABBONE M. & COULET A. (2018). Découverte d'associations entre Événements Indésirables Médicamenteux par les structures de patrons et les ontologies. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 43–48, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- PERSONENI G., DEVIGNES M.-D., DUMONTIER M., SMAIL-TABBONE, MALIKA M. & COULET, ADRIEN A. (2016). Extraction d'associations d'EIM à partir de dossiers patients : expérimentation avec les structures de patrons et les ontologies. In F. MOUGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- PERTIN C., DECCACHE C., GAGNAYRE R. & HAMON T. (2017). Identification du profil des utilisateurs dans les

- forums de discussion de santé. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 85–92, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper18.pdf.
- PETEGROSSO R., PARK S., HWANG T. H. & KUANG R. (2016). Transfer learning across ontologies for phenome–genome association prediction. *Bioinformatics*, p. btw649.
- PINAIRE J., ALOUANE S. B., AZÉ J., BRINGAY S., LANDAIS P. & SALLABERRY A. (2015a). Visualisation interactive de trajectoires de patients. Poster.
- PINAIRE J., AZÉ J., BRINGAY S. & LANDAIS P. (2016). Extraire semi-automatiquement des connaissances dans la littérature biomédicale. In N. PERNELLE, Ed., *27èmes Journées Francophones d'Ingénierie des Connaissances - IC 2016*, Montpellier, France.
- PINAIRE J., AZÉ J., BRINGAY S. & LANDAIS P. (2017a). Infarctus du myocarde : quelles sont les trajectoires de soins pronostiques du décès à l'hôpital? In C. ROUSSEY, Ed., *28èmes Journées Francophones d'Ingénierie des Connaissances - IC 2017*, p. 14–25, Caen, France.
- PINAIRE J., AZÉ J., BRINGAY S., PONCELET P., GENOLINI C. & LANDAIS P. (2017b). Quels événements après un infarctus du myocarde? In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 69–76, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper_4.pdf.
- PINAIRE J., RABATEL J., AZÉ J. & BRINGAY S. (2015b). Recherche et visualisation de trajectoires dans les parcours de soins des patients ayant eu un infarctus du myocarde. In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 5–12, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- PSIUK T. & MANET P. (2015). Intégrer les chemins cliniques dans les outils de soin : rêve ou réalité? In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 13–18, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- RABOUDI A., ALLANIC M., HERVÉ P.-Y., BOUTINAUD P., DURUPT A., BALVAY D. & EYNARD B. (2017). Traçabilité de l'intégration de données biomédicales hétérogènes dans le système SWOMed de gestion du cycle de vie des études biomédicales. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénierie de l'Information Médicale*, p. 57–65, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper_13.pdf.
- RANWEZ S., HARISPE S., CHARLET J., PICARD A.-C. L., JEAN P. & MÉNAGER M.-T. (2015). OntoToxNuc : exploitation d'une ontologie de la toxicologie nucléaire environnementale dans une plateforme collaborative. In N. BRICON-SOUF, Ed., *Actes de SIIM 2015, 3e Symposium sur l'Ingénierie de l'Information Médicale*, p. 19–25, Rennes. <https://www.irit.fr/SIIM/2015/ActesSIIM2015.pdf>.
- REIS J. C. D., PRUSKI C. & SILVEIRA M. D. (2012). Vers une approche automatique pour la maintenance des mappings entre ressources termino- ontologiques du domaine de la santé. In M.-C. JAULENT & L. SOUALMIA, Eds., *Actes de l'Atelier IC pour l'Interopérabilité Sémantique dans les applications en e-Santé*, Paris. <https://sites.google.com/site/icisante2012/programme>.
- RICHARD A., MAYAG B., MEINARD Y., TALBOT F. & TSOUKIAS A. (2018). How AI could help physicians during their medical consultations : An analysis of physicians' decision process to develop efficient decision support systems for medical consultations. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 100–106, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- RICHARD M., AIMÉ X., KREBS M.-O. & CHARLET J. (2013). L'Ingénierie des Connaissances à l'usage du PMSI en Psychiatrie. In L. TAMINE-LECHANI & L. SOUALMIA, Eds., *Actes de SIIM 2013, 2e Symposium sur l'Ingénierie de l'Information Médicale*, Lille. https://www.irit.fr/SIIM/2013/04_siim13.pdf.
- ROSENBLATT F. (1957). The perceptron—a perceiving and recognizing automaton. *Cornell Aeronautical Laboratory*, p. 85–460–1.
- ROUSSEY C. & GHORFI T. A. (2018). Annotation sémantique pour une interrogation experte des Bulletins de Santé du Végétal. In S. RANWEZ, Ed., *29èmes Journées Francophones d'Ingénierie des Connaissances - IC 2018*, p. 37–52, Nancy, France : AFIA.
- RYBARCZYK Y., MARTINS R. & FONSECA J. (2011). Plateforme de traitement multimodal des symptômes aphasiques. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 141–150, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-8.pdf>.
- SCHNELL M., COUFGNAL S., LIEBER J., SALEH S. & JAY N. (2018). Interprétation de bonnes pratiques de codication médicale par du raisonnement à partir de cas — Application à la saisie de données pour les registres du cancer. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 49–54, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- SEDIRI M., MATTA N., LORIETTE S. & HUGEROT A. (2012). Vers une représentation de situations de crise gérées par le SAMU. In S. SZULMAN, Ed., *23èmes Journées Francophones d'Ingénierie des Connaissances - IC 2012*, Paris, France.
- SEGOND F., PONOMAREVA A., RABARIJAONA D., DINI L., KERGOURLAY I., DARMONI S., GICQUEL Q. & METZGER M. H. (2014). Bien représenter pour mieux raisonner : deux approches pour le dossier patient. In S. BRINGAY, N. SOUF & L. TAMINE-LECHANI, Eds., *Actes de l'Atelier IC Santé*, Clermont-Ferrand. <https://www.lirmm.fr/ic-sante/pmwiki/docs/Segond.pdf>.
- SERRES B., ZEMMOURA I., DESTRIEUX C. & VENTURINI G. (2011). Acquisition, visualisation 3d et interactions pour le suivi de dissection. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 103–110, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-9.pdf>.
- SEYMOUR K. & PAYOUX P. (2017). Radiomics Enabler®, an ETL (Extract-Transform-Load) for biomedical imaging in big-data projects. In L. TAMINE-LECHANI, Ed., *Actes de SIIM 2017, 4e Symposium sur l'Ingénie-*

- rie de l'Information Médicale, p. 49–55, Toulouse. https://www.irit.fr/SIIM/2017/SIIM2017_paper_5.pdf.
- SOW A., GUISSÉ A. & NIANG O. (2018). Intégration d'ontologies médicales : amélioration par association des maladies humaines à leurs plus pertinents signes caractéristiques. In S. RANWEZ, Ed., *29èmes Journées Francophones d'Ingénierie des Connaissances - IC 2018*, Nancy, France : AFIA.
- SUN R. (1999). Artificial intelligence : Connectionist and symbolic approaches.
- SÉROUSSI B., GALOPIN A. & GAOUAR M. (2018). Utilisation des cercles thérapeutiques pour l'affichage des recommandations de bonne pratique dans la prise en charge des patients polypathologiques. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 55–61, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- SÉROUSSI B., MESSAI N., LAOUÉNAN C., MENTRÉ F. & BOUAUD J. (2013). Profils patients associés à la non conformité des décisions aux recommandations de prise en charge thérapeutique des cancers du sein : utilisation de l'analyse de concepts formels. In R. TRONCY, Ed., *24èmes Journées Francophones d'Ingénierie des Connaissances - IC 2013*, Lille, France.
- SÉROUSSI B., SAUQUET D., FALCOFF H., JULIEN J. & BOUAUD J. (2011). Formalisation de l'attitude des médecins vis à vis des propositions d'un système d'aide à la décision : évaluation de l'"e-iatrogénie" sur un cas d'hypertension avec ASTI mode guidé. In A. MILLE, Ed., *22èmes Journées Francophones d'Ingénierie des Connaissances - IC 2011*, p. 673–688, Chambéry, France.
- TAPI NZALI M. D., BRINGAY S., LAVERGNE C., OPITZ T., AZÉ J. & MOLLEVI C. (2015). Construction d'un vocabulaire patient/médecin dédié au cancer du sein à partir des médias sociaux. In M.-H. ABEL, Ed., *26èmes Journées Francophones d'Ingénierie des Connaissances - IC 2015*, Rennes, France.
- TERNOIS I., ESCUDIÉ J. B. & DUCLOS C. (2018). Développement et évaluation d'une méthode de codage automatique des endoscopies digestives. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 63–68, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- THEBAULT M., AMARDEILH F., DELAMARRE D., GUILLEMIN-LANNE S., JAMET A. & LILLO-LE LOUET A. (2010). VigiTermes : une plateforme de recherche et d'analyse des publications scientifiques au service de la pharmacovigilance. In S. DESPRÈS, Ed., *21èmes Journées Francophones d'Ingénierie des Connaissances - IC 2010*, p. 183–194, Nîmes, France : Ecole des Mines d'Alès.
- TIRILLY P. (2011). Contributions à l'usage de la modalité dans les systèmes de recherche d'images médicales. In L. TAMINE, S. DARMONI & L. SOUALMIA, Eds., *Actes de SIIM 2011, 1er Symposium sur l'Ingénierie de l'Information Médicale*, p. 87–102, Toulouse. <http://www.irit.fr/SIIM/SIIM2011-8.pdf>.
- TRAORE L., CHNITI A., HUSSAIN S., GRIFFON N., DARMONI S., CHARLET J., SADOU E., OUAGNE D., LEPAGE E. & DANIEL C. (2014). Plateforme d'interopérabilité sémantique gérant les terminologies d'interface au sein d'un espace de partage. In C. FARON-ZUCKER, Ed., *25èmes Journées Francophones d'Ingénierie des Connaissances - IC 2014*, p. 75–86, Clermont-Ferrand, France. Session 1 : Construction, peuplement et exploitation d'ontologies.
- UGON A. (2018). De la difficulté d'évaluer les algorithmes de scorage automatique des données médicales temporelles : exemple de la polysomnographie. In J. CHARLET, M.-D. DEVIGNES & B. SEROUSSI, Eds., *Actes du 3e Atelier IA & Santé*, p. 69–74, Nancy. http://pfia2018.loria.fr/communicationsiasante_3juillet2018/.
- UGON A., KOTTI A., SEDKI K., PHILIPPE C., SÉROUSSI B., BOUAUD J., GANASCIA J.-G., GARDA P. & PINNA A. (2016). Reconnaissance des stades de sommeil à l'aide d'un outil de support à la décision basé sur les connaissances et la pratique des experts. In F. MOUNGIN, A. ABDAOUI & P. ZWEIGENBAUM, Eds., *Actes du 2e Atelier IA & Santé*, Montpellier. https://ic2016.sciencesconf.org/conference/ic2016/pages/IA_Sante.pdf.
- WANG H., DING Y., TANG J., DONG X., HE B. & QIU J. (2011). Finding complex biological relationships in recent pubmed articles using bio-lda. *PLoS One*, **6**(3), 69.
- WANG S.-H., DING Y., ZHAO W., HUANG Y.-H., PERKINS R., ZOU W. & CHEN J. J. (2016). Text mining for identifying topics in the literatures about adolescent substance use and depression. *BMC Public Health*, **16**(1), 279.
- ZNAIDI E., TAMINE L., CHOUQUET C. & LATIRI C. (2013). Analyse exploratoire des requêtes d'experts médicaux : cas des campagnes d'évaluation TREC et CLEF. In L. TAMINE-LECHANI & L. SOUALMIA, Eds., *Actes de SIIM 2013, 2e Symposium sur l'Ingénierie de l'Information Médicale*, Lille. https://www.irit.fr/SIIM/2013/02_siim13.pdf.