

Un modèle de données adapté à la recherche d'information dans le dossier patient informatisé : étude, conception et évaluation

Ahmed-Diouf Dirieh Dibad¹, Lina F. Soualmia^{1,2}, Tayeb Merabti¹, Julien Grosjean¹, Saoussen Sakji¹, Philippe Massari¹, Stéfan J. Darmoni¹

¹*CISMéF, CHU de Rouen & TIBS, LITIS EA 4108, Université de Rouen, France*

²*LIM&Bio, EA 3969, Université Paris 13, Sorbonne Paris Cité, France*

Abstract

Background : Electronic Health Records (EHRs) have an important role in medical decision making and are considered as being a major part of medical support system. To ensure this role effectively, the information must be easily found, even in voluminous and large number of EHRs. This requires to develop search capabilities for EHR information retrieval (IR). Methods : To perform this, we propose a data model adapted to IR. The data analysis of EHRs of the Rouen University Hospital led us to consider them as being a set of informational elements linked by temporal and conceptual relationships. This model consists of two main entities to manage the information elements and their metadata. After implementation, we have evaluated the capacity of the proposed model to take into account all data from the EHRs and its adaptation to IR. Results : We have performed 25 queries on EHR (20 correspond to single patient and 5 to multi-patients). The results in 19 cases (including 3 incomplete results) were evaluated as successful. These results confirm the ability of the model to take into account most of relevant data of EHR in IR and clearly show the differences in complexity of queries between the initial and our model of IR. Conclusion : The preliminary evaluation of the data model we have developed has shown its adaptation to IR in EHR. Nevertheless, further work on larger sets is needed to confirm this.

Keywords

Decision Support Techniques; Electronic Health Records; Information Storage and Retrieval; Vocabulary, controlled.

1 Introduction

Le dossier patient informatisé (DPI) [1], du fait des progrès technologiques prend une place croissante dans notre système de santé, sous forme de dossiers partagés au sein d'un établissement ou de réseaux de soin et dans un futur proche par la mise en place du Dossier Médical Personnel (DMP). Les applications de gestion des dossiers patients permettent l'enregistrement des éléments constituant le dossier médical et la communication entre les professionnels partageant la prise en charge des patients. Le DPI est devenu un élément important dans la prise de décision médicale et est considéré comme faisant partie des systèmes d'aide à la décision médicale [2].

Outre ces fonctions d'enregistrement et de communication, l'un des apports importants des

technologies de l'information et de la communication (TIC) est la possibilité de retrouver plus facilement l'information [3]. Les applications de gestion actuelles ne proposent pas de fonctions de recherche d'information dans les dossiers patients informatisés (RI) alors qu'elles deviennent indispensables du fait de l'augmentation du nombre d'éléments contenus dans chaque dossier [4,5].

Les informations du DPI sont pour une faible partie structurées (données démographiques, de prises en charge hospitalières, codage des diagnostics et des actes, résultats d'analyses biologiques) et pour l'essentiel sous forme textuelle, et donc non structurées, dans les courriers et comptes-rendus, qu'ils soient d'hospitalisation ou d'actes. Développer des outils de RI impose donc d'indexer ces dernières à l'aide des terminologies de référence grâce à des outils de Traitement du Langage Naturel (TAL) et/ou des outils d'indexation sémantique fondés sur les principales terminologies et ontologies de santé.

Disposer de données structurées et indexées est une condition préalable au développement de fonctions de RI, mais ce n'est pas suffisant, ces données devront être enregistrées dans une base de données possédant un modèle adapté à la RI, et indépendant du modèle de données du système de gestion de DPI.

La conception de ce modèle adapté à la recherche d'information et permettant d'utiliser le langage d'interrogation SPARQL¹ est l'objet du travail que nous présentons ici.

2 Matériel et méthodes

Afin de concevoir le modèle, de tester son adaptation à la prise en compte des informations du DPI et à la RI, nous nous sommes appuyés sur le DPI existant au CHU de Rouen. L'informatisation du dossier patient du CHU [6] de Rouen a débuté en 1992, depuis cette date il prend en compte les données administratives et démographiques, les données cliniques et biologiques (résultats d'analyses biologiques, actes médicaux, comptes-rendus de séjours, courriers, etc.) dans le cadre d'un dossier partagé.

Le DPI est actuellement géré par l'application C Page Dossier Patient (CDP) qui contient plus de 4 millions de comptes-rendus et d'autres courriers pour les 800.000 dossiers informatisés. Les données du DPI sont stockées dans une base Oracle, sous forme structurée pour certaines données (données démographiques, prise en charge hospitalières, codage des diagnostics et des actes, résultats d'analyses biologiques) et en format texte pour l'ensemble des courriers et comptes-rendus médicaux, une grande partie des informations ne figure que dans ces textes. La base de données du système CDP contient de nombreuses tables (une cinquantaine) gérant les données du DPI, ce qui nécessiterait le développement de requêtes spécifiques pour la RI sur les données structurées. Rechercher une information médicale contenue à la fois dans les données codées et les données textuelles, est difficile tant au niveau de la rapidité de réponse qu'à l'accès à l'information elle-même, dans ce type de modèle de données. Les modèles proposés [7,8,9,10] dans la littérature posent les mêmes problèmes pour la RI.

Ceci nous a amené à concevoir un modèle de données indépendant du modèle de données du système de gestion des dossiers pour effectuer les RI.

Dans un premier temps, nous avons analysé les données structurées contenues dans CDP, afin de lister les informations qu'elles contiennent, de préciser leurs formats et de projeter leurs modes de transposition. Le modèle à concevoir doit pouvoir gérer ces données, mais également les informations issues des courriers après indexation. L'ensemble des données doit pouvoir être pris en compte sous forme de métadonnées s'appuyant sur les

¹ <http://www.w3.org/TR/rdf-sparql-query/>

terminologies CIM10, CCAM, LOINC, SNOMED, ATC qui font partie de l'univers multi-terminologique sur lequel s'appuieront les outils de RI [11].

Outre la prise en compte de ces métadonnées, le modèle cible devra être constitué d'un nombre de tables limité afin de pouvoir utiliser un moteur de recherche terminologique ou le SPARQL avec des temps de réponses minimums (de l'ordre de la seconde pour une requête sur un patient unique et de la minute pour une requête épidémiologique sur plusieurs patients) : c'est l'étape d'optimisation qui est essentielle à l'utilisation future du moteur de recherche. En prenant en compte toutes ces contraintes, un modèle de données a été défini et est représenté dans la figure 2. A titre expérimental, il a été implémenté sous Oracle. Les terminologies, issues du système d'information de CISMef, sont transformées et stockées dans une base sémantique sous forme de triplets RDF. Ces derniers consistent en des associations de sujet, prédicat et objet. L'exécution des requêtes sur les triplets RDF est réalisée en utilisant le protocole SPARQL.

Notre travail, en plus de la conception du modèle, nous a permis d'évaluer sa capacité à prendre en compte les données structurées (données démographiques, de prises en charge hospitalières, codage des diagnostics et des actes, résultats d'analyses biologiques) de 20 dossiers médicaux anonymisés provenant de CDP. Une évaluation préliminaire de son adaptation à la RI a été effectuée, sur 20 cas tests pour la RI concernant un patient et sur 5 pour la RI multi-patients. Ces cas tests ont été créés par un des auteurs (PM) au vu des dossiers originaux de CDP. Ces cas tests incluent plusieurs critères de recherche, un court résumé du dossier ainsi que les résultats attendus.

2.1 Présentation du modèle

2.1.1 Principes de base

L'analyse des données nous a permis dans un premier temps de définir un sous-modèle au sein du modèle physique de CDP et de constater que le DPI pouvait être considéré comme un document contenant des éléments informationnels liés par des relations conceptuelles et temporelles. Un élément informationnel peut être un patient donné, un séjour d'hospitalisation pour une fracture de jambe, une radiographie de la cheville, un résultat d'analyse de l'hémoglobine ou une information indexée provenant d'un compte-rendu de séjour.

Le modèle que nous proposons s'abstrait de la vision générale dans laquelle on perçoit un dossier médical comme un ensemble d'évènements médicaux [12]. Nous considérons ici le DPI comme étant un ensemble d'éléments informationnels qui sont décrits par des métadonnées et des vocabulaires, de manière similaire à une ressource Web [13] comme illustré dans le tableau 1. Ainsi un séjour d'hospitalisation sera un élément informationnel, codé par un ou plusieurs codes CIM-10 pour les pathologies diagnostiquées et caractérisé par une ou plusieurs métadonnées.

Tableau 1 : Comparaison d'indexation d'une donnée clinique et d'une ressource Web

| | Dossier Patient Informatisé | Ressource Web |
|--------------------|---|---|
| Information | IPECUF_RM (N°Séjour) : 114005 Date Entrée : 06/04/07 Date Sortie : 18/05/07 Modalité Entrée : Domicile Modalité Sortie : Domicile Statut Séjour : Hospitalisation Diagnostic : épilepsie | Titre : "Asthme et piscines", 2009 URL : http://www.swiss-paediatrics.org/sites/default/files/paediatrica/vol20/n4/pdf/50-53.pdf |
| Indexation | CIM-10 : G40 | MeSH : D0-01249 |
| Métadonnées | Type d'information : Séjour Type Séjour : Hospitalisation Modalité Séjour : Domicile Période : 06/04/07 au 18/05/07 | Type Ressource : article périodique Auteur : Société Suisse de Pédiatrie Année : 2009 |

2.1.2 Description du modèle

Le modèle est schématisé dans la figure 1. L'entité «**Element_Information**» constitue l'entité de base. Cette entité peut prendre en compte toutes des données du DPI, qu'elles soient des données démographiques, de prises en charge, des codes diagnostiques, des codes d'actes, des données structurées, ou encore des métadonnées issues de l'indexation des courriers.

L'entité «**Element_RelationInformation**» gère les relations conceptuelles et temporelles entre deux éléments informationnels et permet de garder la structure logique du DPI, liaison d'un séjour S au patient A par exemple.

Les autres entités «**Element_TypeInformation**», «**Element_TypeAttribut**», «**Element_TypeRelation**» constituent les métadonnées de notre modèle. L'entité «**Element_RelationDescripteur**» permet de gérer l'indexation d'un élément informationnel par un ou plusieurs descripteurs de nos terminologies médicales. Cette indexation est issue soit des données structurées et correspond au codage PMSI², soit issue de l'indexation automatique du compte-rendu en format texte à l'aide de l'outil F-MTI [14].

A titre d'exemple, pour le séjour cité dans le tableau 1, l'entité «**Element_Information**» contiendra l'enregistrement suivant : 'SEJ_114005_1' comme *ID_Information*, 'SEJOUR' pour *ID_TypeInformation*, '114005' pour *ID_Origine*, '06/04/07' pour *DateDebut* et '18/05/07' pour *DateFin*. Les attributs spécifiques sont décrits par l'entité «**Element_Attribut**». Dans ce cas *ID_TypeAttribut* sera 'TYPE_SEJOUR', et *ValeurAttribut* sera 'Hospitalisation'.

² Programme de Médicalisation des Systèmes d'Information

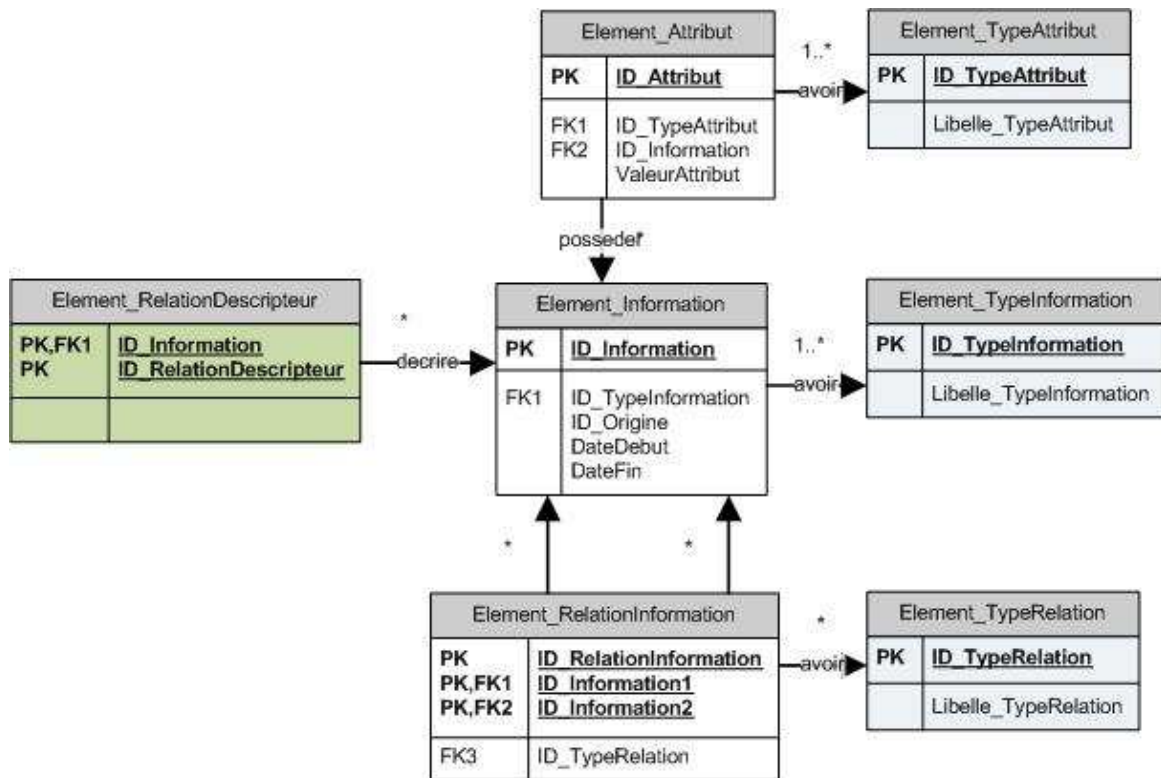


Figure 1 : Le modèle de représentation d'un dossier patient informatisé

2.2 Processus de Recherche d'Information au sein du DPI

La base de données correspondant au modèle proposé est l'élément central de la RI dans le DPI telle que décrite dans la figure 2. Grâce à des programmes de réplication adaptés (de type Extract Transform Load (ETL)), les données du DPI sont intégrées dans notre modèle. Le premier de ces outils convertit les données dans des formats de transition standards conformes au standard Health Level Standard Seven (à HL7³) via une Application Programming Interface (API).

3 Résultats

3.1 Capacité du modèle à prendre en compte les données du DPI

L'ensemble des données structurées de 20 dossiers anonymisés, a été répliquées dans la base, au format défini dans le modèle incluant les attributs généraux et ceux spécifiques à chaque type de données. Les données ayant été répliquées correspondaient :

- aux données démographiques (Nom, Prénom, date de naissance, sexe, etc.)
- aux prises en charge (hospitalisations, consultations)
- aux actes
- aux codages diagnostiques (CIM10)
- aux codages des actes (CCAM)

³ <http://www.hl7.org/>

Le modèle que nous proposons s'avère être adapté à la prise en compte de l'ensemble des données.

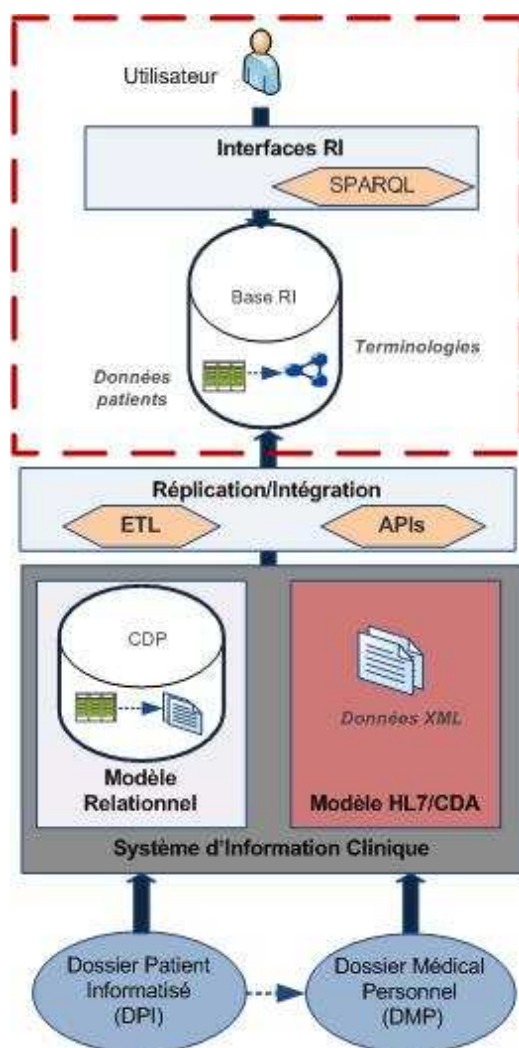


Figure 2 : Schéma global du processus de RI proposé

3.2 Comparaison des requêtes faites sur la base DPI et sur la base correspondant au modèle

La comparaison des requêtes a été effectuée sur des exemples de RI de complexité intermédiaire.

- Recherche des séjours liés à une pathologie infectieuse chez un patient donné. Les codes CIM10 correspondant à une pathologie infectieuse n'ont pu être déterminés directement et nous ont obligés à utiliser le métaterme "infectiologie"(un super-concept pour agréger des concepts de même spécialité médicale [15]).
- Recherche d'un acte radiologique : comme par exemple la radiographie du bassin chez un patient donné.

Les requêtes sur notre base, conforme au modèle, n'incluent que 2 tables et une jointure commune aux 2 requêtes. Les requêtes dans la base DPI contiennent 4 tables et 3 jointures pour chacune d'entre elles, mais 2 des tables et 2 des jointures sont différentes pour ces requêtes. Des exemples comparant les requêtes SQL faites sur la base DPI et celles en SPARQL sur une base conforme au modèle figurent dans le tableau 2.

Tableau 2 : Comparaison d'une requête SQL et SPARQL pour un test case

| Test case | Requête sur la base DPI | Requête sur le modèle |
|---|---|--|
| TC n°1 : "Rechercher les épisodes infectieuses du patient n°1" | <pre>SELECT PATIENT.IDPATIENT, METATERMES_CIM.LIBELLE, PEC.UM_SEJ, PEC.DATEENTREE, PEC.DATESORTIE, DIAG.CODECIM10, DIAG.LIBELLEDIAG, METATERMES_CIM.METAT FROM (METATERMES_CIM INNER JOIN DIAG ON METATERMES_CIM.CODE = DIAG.CODECIM10) INNER JOIN (PATIENT INNER JOIN PEC ON PATIENT.IDPATIENT = PEC.IDPATIENT) ON DIAG.IDPECUF_RM = PEC.IDPECUF_RM WHERE (((PATIENT.IDPATIENT)=1) AND ((METATERMES_CIM.METAT)="infectiologie"));</pre> | <pre>SELECT d.ID_INFORMATION, r.ID_INFORMATION1, DATE_RELATION FROM ELEMENT_RelationDescripteur d LEFT OUTER JOIN ELEMENT_RelationInformation r ON d.ID_INFORMATION = r.ID_INFORMATION2 WHERE r.ID_TypeRelation= 'T_REL_SEJ_PAT' and r.ID_INFORMATION1= 'PATIENT1' and ID_DESC_URI IN (SELECT c2 FROM TABLE (sem_match(' (?r :type <publishing#BT-NT>) (?r <publishing#BT> ?c1) (?r <publishing#NT> ?c2) (?c1 <smts#CIM10MT> "infectiologie ") ', sem_models('CIM10'), null, null, null)))</pre> |
| TC n°2 : "Rechercher une radiographie du bassin du patient n°6" | <pre>SELECT PATIENT.IDPATIENT, PEC.UM_SEJ, PEC.DATEENTREE, PEC.DATESORTIE, ACTE.UF_ACTE, ACTE.DATEACTE, CCAM.LIBELLECODE FROM CCAM INNER JOIN (ACTE INNER JOIN (PATIENT INNER JOIN PEC ON PATIENT.IDPATIENT = PEC.IDPATIENT) ON ACTE.IDPECUF_RM = PEC.IDPECUF_RM) ON CCAM.IDACTE = ACTE.IDACTE WHERE (((PATIENT.IDPATIENT)=1) AND ((CCAM.["LIBELLECODE"]) Like "bassin*"));</pre> | <pre>SELECT d.ID_INFORMATION, r.ID_INFORMATION1, DATE_RELATION FROM ELEMENT_RelationDescripteur d LEFT OUTER JOIN ELEMENT_RelationInformation r ON d.ID_INFORMATION = r.ID_INFORMATION2 WHERE r.ID_TypeRelation = 'T_REL_ACT_PAT' and r.ID_INFORMATION1= 'PATIENT6' and ID_DESC_URI IN (SELECT c2 FROM TABLE (sem_match(' (?c2 <rdf- schema#label> ?label)', sem_models('DM_CCAM'), null, null, null)) WHERE REGEXP_LIKE (LOWER(label), '(.)*(bassin)(.)*'))</pre> |

3.3 Evaluation de l'adaptation du modèle à la RI

3.3.1 Recherche d'information dans le dossier d'un patient

Elle a été effectuée à partir de 20 cas tests de RI (sous forme texte) : 9 correspondaient à des recherches d'un ou plusieurs séjours liés à un diagnostic donné, 6 à la recherche d'un ou plusieurs actes, 2 à des prises en charge dans un service donné et 3 à la survenue de types de pathologies (accidents médicamenteux, épisodes infectieux).

Tableau 3 : Evaluation des résultats RI mono patient

| Echelle / Type de requête | n | Conformes aux résultats attendus | Incomplets par rapport aux résultats attendus | Ne correspondant pas aux résultats attendus |
|------------------------------|---|----------------------------------|---|---|
| Diagnostique | 9 | 7 | 2 | |
| Sur des actes | 6 | 5 | 1 | |
| Sur des prises en charge | 2 | | | 2 |
| Sur des types de pathologies | 3 | 1 | | 2 |

Les résultats de l'évaluation figurent dans le tableau 3, dans seulement 4 cas la recherche ne retrouve pas les résultats attendus, dans 2 cas il s'agit de recherche de prises en charge. La non réponse est expliquée par l'absence de l'information (libellés des services de séjours) dans la base, dans les 2 autres cas seules des stratégies de recherche (actuellement non implémentées) auraient pu retrouver les informations.

3.3.2 Recherche d'information multi-patients

La même évaluation a été réalisée sur plusieurs dossiers patients : 2 recherches incluant des données démographiques et des pathologies, 3 sur la recherche de types de pathologies (pathologie infectieuse, accidents médicamenteux) ou des types d'actes (intervention chirurgicale).

Tableau 4 : Evaluation des résultats RI multi patient

| Echelle / Type de requête | n | Conformes aux résultats attendus | Incomplets par rapport aux résultats attendus | Ne correspondant pas aux résultats attendus |
|---|---|----------------------------------|---|---|
| Sur des pathologies et des données démographiques | 2 | 2 | | |
| Sur des types de pathologies ou d'actes | 3 | 2 | | 1 |

Les résultats de l'évaluation figurent dans le tableau 4. Seule la recherche d'accidents médicamenteux n'a pas fourni de réponse, mais ceci n'était pas lié au modèle mais à l'absence actuelle de ce concept dans la terminologie. La recherche d'interventions chirurgicales n'a été possible qu'en utilisant une stratégie de recherche reposant sur les caractéristiques des codes CCAM.

4 Discussion

L'informatisation du dossier médical n'a pas pour but de supprimer le dossier papier mais a pour objectif de rendre les informations importantes accessibles et utilisables [16]. Pour qu'elles soient utilisables elles doivent être facilement retrouvées au mieux par des fonctions de recherche d'information qui font actuellement défaut. En effet, Arnott Smith [17] en 2008 ne retrouve dans la littérature que 2 articles centrés sur la RI dans le DPI. Plus récemment, les travaux de Dinh [18,19] et de Currie [20] concernent cette problématique.

Pour Rector [21], la recherche d'information pertinente dépend de la structure utilisée pour représenter un DPI. Ceci a conduit plusieurs auteurs à proposer des modèles de structuration du contenu du DPI, adaptés à la RI. Le modèle de Rogers [7] qui représente l'histoire médicale et chronologique du patient comme un réseau sémantique de concepts médicaux ou encore le modèle de Lindemann [8] qui a transformé le dossier médical sous forme de triplets RDF ou encore sous la forme de documents XML pour Pinon [9]. D'autres travaux ont porté sur la projection des événements cliniques du DPI sur des axes de vocabulaire utilisés comme critères de recherche [22].

Les techniques de RI utilisables sont dépendantes du modèle. Par exemple pour le modèle RDF [8], l'utilisation du SPARQL permet d'optimiser les recherches. Pour le modèle CLEF [7], les recherches sont dépendantes du formalisme du modèle.

Dans le modèle de données que nous proposons, la RI peut être réalisée via des requêtes types dont le seul élément variable est la partie terminologique de la clause « *where* ». La première évaluation que nous avons effectuée confirme son aptitude à gérer les données du DPI, et son adaptation à la RI tant au sein d'un seul dossier que dans une base multi-patients. Ce modèle permettra notamment de gérer les données biologiques (résultats d'analyses biologiques, etc.) et les données (métadonnées, valeurs numériques et symboliques, etc.) provenant de l'indexation multi-terminologique des comptes-rendus, qui ont le même format que les données codées du DPI. Et de ce fait, il constituera un modèle d'intégration des données structurées et textes comme [10] à l'instar des solutions d'intégration existantes : utilisation d'ontologie pour Diallo [23], utilisation des outils du Web Sémantique pour Sahoo [24].

La structure des données que nous avons analysée, celle du DPI du CHU de Rouen, respecte la norme HISA [25] ce qui doit permettre sa transposition à d'autres systèmes de gestion de DPI respectant cette norme. La réplique des données via un passage au format HL7 devrait pouvoir en étendre l'utilisation.

L'évaluation actuelle du modèle dans la RI est limitée : comme pour de nombreux DPI la quantité de données textuelles est la plus importante [26]. La prise en compte après indexation des informations contenues dans les textes et des résultats d'analyses biologiques devra faire l'objet de travaux ultérieurs. Ces travaux devront confirmer que le modèle est adapté aux différents besoins de recherche d'information. En particulier ceux précédemment étudiés, recherche d'information contenue dans des comptes-rendus [27,28] ou recherche spécifique pour l'évaluation de la qualité du soin du patient [29], pour des essais cliniques [30] ou encore pour la recherche épidémiologique [31].

La capacité de ce modèle à fournir les données nécessaires à la présentation du dossier sous forme de vues "orienté temps", "orienté problème" [32], par spécialité [33] ou sous forme d'une synthèse [34] devra également être évaluée. Une de ses limites pourrait être la gestion temporelle des informations dans des vues chronologiques ou lors de la recherche par exemple d'infections étant apparues après une intervention.

5 Conclusion

Nous avons décrit dans cet article un modèle de données pour représenter le DPI dans le but de faciliter la recherche d'information. L'évaluation préliminaire a montré son adaptation à la RI. En revanche, les travaux futurs que nous envisageons devront le confirmer et permettront d'en préciser les éventuelles limites qui n'auraient pas été identifiées au stade préliminaire.

Références

- [1] In: *Informatique Médicale*. Degoulet P, Fieschi M, eds. Paris: Abrégés Masson 1998; pp. 120-30.
- [2] Une étude sur les systèmes d'aide à la décision médicale. Rapport remis à la Haute Autorité de Santé, juin 2010. Disponible sur : <http://www.has-sante.fr/portail/jcms/c_989360/systemes-informatiques-d-aide-a-la-decision-medicale> (Consulté le 20.10.2010).
- [3] Grossman DA, Frieder O, Holmes DO, Roberts DC. Integrating Structured Data and Text: A Relational Approach. *Journal of the American Society of Information Science* 1997; 48(2): 122-32.
- [4] Laerum H, Ellingsen G, Faxvaag A. Doctors' use of electronic medical records systems in hospitals: Cross sectional survey. *BMJ* 2001; 323(7325): 1344-8.
- [5] Gorman PN, Helfand M. Information Seeking in Primary Care: How Physicians Choose Which Clinical Questions to Pursue and Which to Leave Unanswered. *Medical Decision Making* 1995; 15: 113-9.
- [6] Massari P, Smuraga I, Froment L, Boudehent S, Czernichow P, Streiff J, Baldenweck M, Hecketsweiler P. Application de gestion les dossiers informatisé du CHU de Rouen. *Cinquièmes Journées Francophone d'Informatique Médicale* 1994; pp. 321-20.
- [7] Rogers J, Puleston C, Rector A. The CLEF Chronicle: Patient Histories Derived from Electronic Health Records. *The 22nd International Conference on Data Engineering Workshops (ICDEW'06)* 2006; 109.
- [8] Lindemann G, Schmidt D, Schrader T, Keune D. The Resource Description Framework (RDF) as a Modern Structure for Medical Data. *World Academy of Science, Engineering and Technology* 2007; 240-43.
- [9] Pinon JM, Calabretto S, Pouillet L. Document Semantic Model: An Experiment with Patient Medical Record. In: *Proceedings of ICC/IFIP Conference on Electronic Publishing: EP'97, New Models and Opportunities*, 1997; 262-72
- [10] Branson A, Hauer T, McClatchey R, Rogulin D and Shamdasani J. A Data Model for Integrating Heterogeneous Medical Data in the Health-e-Child Project. *Stud Health Technol Inform* 2008; 138: 13-23.
- [11] Joubert M, Merabti T, Vandebussche PY, Abdoune H, Dahamna B, Fieschi M, Darmoni SJ. Modeling and Integrating Terminologies into a French Multi-terminology Server. In: *MEDINFO* 2010 (in press).
- [12] Huff SM, Rocha RA, Bray BE, Warner HR, Haug PJ. An event model of medical information representation. *J Am Med Inform Assoc* 1995; 2(2): 116-34.
- [13] Dirieh Dibad AD, Sakji S, Prieur E, Pereira S, Joubert M, Darmoni SJ. Recherche d'information multiterminologique en contexte : étude préliminaire. *Informatique et Santé* 2009; 17: 101-12.

- [14] Pereira S. Indexation multi-terminologique de concepts en santé. Thèse de l'Université de Rouen. 2008.
- [15] Gehanno JF, Thirion B, Darmoni SJ. Evaluation of meta-concepts for information retrieval in a quality-controlled Health Gateway. *AMIA Annu Symp Proc* 2007; 11: 269-73.
- [16] Ondo K, Wagner J, Gale K. The electronic medical record: Hype or reality? *Journal of Healthcare Information Management* 2002; 17(4): 2.
- [17] Smith A. The electronic medical record as a new domain. *Proceedings of the American Society for Information Science and Technology* 2006; 43: 1-30.
- [18] Dinh D, Tamine L. Vers un modèle d'indexation sémantique adapté aux dossiers médicaux de patients. *CORIA* 2010; 325-36.
- [19] Dinh D, Tamine L. Recherche d'Information dans les documents biomédicaux : Approche basé sur le sens précis des concepts. *INFORSID* 2010; 261-74.
- [20] Currie AM, Cohan J, Zlatic L. Information Retrieval of Electronic Medical Records. *CICLing* 2001; 2004: 460-71.
- [21] Rector AL, Nowlan WA, Kay S. Foundations for an electronic medical record. *Methods Inf Med* 1991; 30: 179-86.
- [22] Corley JH, Underwood WE. Scotch and Sodas: Semantically Oriented Data Access and Storage for a Medical Record DBMS. *Proc Annu Symp Comput Appl Med Care* 1979; 17: 98-104.
- [23] Diallo G, Simonet M, Simonet A. Bringing Together Structured and Unstructured Sources: the OUMSUIS Approach. *Proceedings of OTM Confederated International Workshops and Posters, Knowledge Systems in Bioinformatics (KSinBit'06) in conjunction with OTM'06* 2006. LNCS 4277: 699-709.
- [24] Sahoo S, Bodenreider B, Rutter J, Skinner K, Sheth A. An ontology-driven semantic mashup of gene and biological pathway information: Application to the domain of nicotine dependence. *Journal of Biomedical Informatics* 2008; 41: 752-65.
- [25] Health Informatics Service Architecture (HISA), An informal guide, septembre 2009 [Accesible à l'URL : <http://www.consortioedith.it/public/HISA%20-%20InformalGuideto%20its%20role-v1-3.pdf> , visualisé le 20 octobre 2010].
- [26] Rector AL. Clinical terminology: why is it so hard?. *Methods Inf Med* 1999; 38: 239-52.
- [27] Spat S, Cadonna B, Rakovac I, Gütl C, Leitner H, Stark G, Beck P. Enhanced Information Retrieval from Narrative German-Language Clinical Text Documents Using Automated Document Classification. *Studies In Health Technology And Informatics* 2008; 136: 473-78.
- [28] Chou S, Chang W, Cheng C, Jehun JC, Chang C. Information Retrieval System for Medical Records and Documents. *The 30th Annual International IEEE EMBS Conference* 2008; 1474-7.
- [29] Brown PJB, Sonksen P. Evaluation of the quality of information retrieval of clinical findings from a computerized patient database using a semantic terminological model. *J Am Med Inform Assoc* 2000; 7(4): 392-403.
- [30] Patel C, Cimino JJ, Dolby J, Fokoue A, Kalyanpur A, Kershenbaum A, Ma L, Schonberg E, Srinivas K. Matching patient records to clinical trials using ontologies. *Proc ISWC07* 2007. LNCS 4825: 816-29.
- [31] Powell J, Buchan I. Electronic health records should support clinical research. *J Med Internet Res* 2005; 7(1): e4.

- [32] Bainbridge M, Salmon P, Rappaport A, Hayes G, Williams J, Teasdale S. The problem Oriented Medical Record - just a little more structure to help the world go round? Clinical Computing Special Interest Group (CLICSIG) of the PHCSG. *Proceedings of the 1996 Annual Conference of The Primary Health Care Specialist Group of the British Computer Society* 1996.
- [33] Massari P, Pereira S, Thirion B, Derville A, Darmoni SJ. Use of Super-Concepts to customize electronic medical records data display. *Stud Health Technol Inform* 2008; 136: 845-50.
- [34] Afantenos SD, Karkaletsis V, Stamatopoulos P. Summarization from medical documents: a survey. *Artificial Intelligence in Medicine* 2005; 33(2): 157-77.

Adresse de correspondance

Stéfan Darmoni, Equipe CISMéF, Cour Leschevin, Porte 21, 3^{ème} étage, 1 rue de Germont
76031 Rouen Cedex;

Courriel : stefan.darmoni@chu-rouen.fr