

SIMS@REIN : un système d'information multi-sources pour l'insuffisance rénale terminale

Paul Landais^{a*}, Ana Simonet^b, Didier Guillon^b, Christian Jacquelinet^c, Mohamed Ben Saïd^a, Claude Mugnier^a, Michel Simonet^b

^a *Laboratoire de biostatistique et d'informatique médicale (LBIM), université Paris-5, UFR Necker–Enfants-Malades, hôpital Necker, 149, rue de Sèvres, 75743 Paris cedex 15, France*

^b *Laboratoire des techniques de l'imagerie, de la modélisation et de la cognition (UJF–TIMC), CNRS UMR 5525, université Joseph-Fourier, BP 53X, 38041 Grenoble cedex, France*

^c *Établissement français des greffes, 5, rue Lacuée, 75012 Paris, France*

Reçu le 18 juillet 2001 ; accepté le 23 novembre 2001

Présenté par Michel Thellier

Abstract – SIMS@REIN: a multi-source information system for end-stage renal disease. In France, the prevalence of End-Stage Renal Disease (ESRD) is not precisely known. The sources of information are scattered and not coordinated. Consequently, care is ill adapted to meet the demand. The Multi-Source Information System is the basis of the Renal Epidemiology and Information Network (REIN). It is dedicated to improve and organise our medical and epidemiological knowledge of ESRD and to aid public health decision-making in this area. The proposed approach is based on the datawarehouses. This model allows a unified vision of scattered data into distinct databases, for a better management, be it particular (patient follow-up) or global (regional follow-up), with a finality of aid in decision-making. Several categories of problems were considered: the global conception of the information system, the organisation of the datawarehouse, which offers different viewpoints of the data, the integration of heterogeneous data coming from different sources, data exchange and definition of a specific ontology. **To cite this article:** P. Landais et al., *C. R. Biologies 325 (2002) 515–528*. © 2002 Académie des sciences / Éditions scientifiques et médicales Elsevier SAS

multi-sources information system / datawarehouse / n-tier architecture / end-stage renal disease

Résumé – La prévalence de l'insuffisance rénale terminale (IRT) n'est pas connue de façon précise. Les sources d'informations sont parcellaires ou catégorielles et non coordonnées. En conséquence, l'offre de soins est inadaptée à la demande. La mise en place du système d'information multi-sources (SIMS), qui constitue la base du réseau épidémiologie et information en néphrologie (REIN), vise à améliorer notre connaissance médicale, épidémiologique et organisationnelle, pour la prise de décisions de santé publique pour l'IRT. L'approche proposée se fonde sur les entrepôts de données. Ce modèle permet, à partir des faits dispersés dans différentes bases, d'obtenir une vision unifiée des données à des fins de meilleure appréhension de la situation, à la fois particulière (suivi du patient) et globale (suivi régional), avec une finalité d'aide à la décision sanitaire. Plusieurs catégories de problèmes ont été abordées : la conception globale du système, l'organisation de l'entrepôt, l'intégration de données hétérogènes provenant de sources diverses, l'échange des données et la prise en compte d'une ontologie spécifique. **Pour citer cet article :** P. Landais et al., *C. R. Biologies 325 (2002) 515–528*. © 2002 Académie des sciences / Éditions scientifiques et médicales Elsevier SAS

système d'information multi-sources / entrepôt de données / architecture n-tiers / insuffisance rénale terminale

*Correspondance et tirés à part.

Adresse e-mail : landais@necker.fr (P. Landais).

Abridged version

1. Introduction

The goal of the SIMS@REIN system is to constitute the informational support of the Renal Epidemiology and Information Network (REIN), which targets to coordinate the interoperability of the participants and of the information systems with the establishment of a Multi-Source Information System (MSIS). This system concerns both dialysed and transplanted patients. It depends on a continual and exhaustive recording of new cases and their follow-ups. It collects, in a standardised manner, a minimal set of data, the characteristics of which have been defined with professionals. It exchanges exhaustive and validated data with the respect of data security and of the right for patient access, based on the principle of a secured Intranet.

2. Interoperability elements

2.1. The report

The computerisation of the Nephrology teams is heterogeneous. Technically, there are two solutions for creating a new information system: either developing ad hoc programs as one goes along, bit by bit, depending on the needs, or to put in place a Multi-Source Information System that houses the computer applications and existing data but that expands their feasibility. A Multi-Sources Information System (MSIS) depends on a number of sources, distributed and heterogeneous, which propose a subset of their data to job-sharing, possibly conditioned by rights to access. For the system to be able to respond to a combination of queries specific to problems from its origin, it must stay open to allow for the integration of new components or to respond to new demands. The datawarehouses, the interoperable engines, the Web applications on Intranet or Internet are all examples of MSIS. Once the users have a MSIS at their disposal, they only have to be trained for one system, the movement between the engines being transparent.

One of the largest difficulties of the MSIS's implementation originates from the heterogeneity of the component databases at different levels: the *system* level, characterised by differences in operating systems (Unix, NT...), in protocols (TCP/IP, IPX...) or in systems for management of databases (Oracle, Access, Omnis, files...), and the *semantic* level, where one refers to the different ways in which the real world entities are modelled and represented. The heterogeneity of the systems is managed by *middleware* tools and by the expansions brought to their system by the SGBD designers.

The interoperability of the information systems can be defined as the capacity of multiple systems of information to not only exchange information (functional interoperability) but also to utilise exchanged information (semantic interoperability).

2.2. Interoperability at the semantic level

One can consider four large families of software architectures to resolve the semantic heterogeneity, the translator approach, the integrated approach, the completely decentralised approach and the middleware approach. The latter is based on a middleware that has a certain number of services at its disposal. These services particularly use a conflict detector, an ontology of domain and a library of local meta-knowledge descriptions, also known as local contexts. One advantage to this solution is that the semantic heterogeneity solution is rendered invisible for the users, because it continues to use the user's own language. The counterpart, however, is that this solution emphasises the putting in place of services for the running operation of information (the middleware services).

The global ontology must be defined. It is based on the core, previously defined, and will also need to dispose of necessary information for exchanges between local databases, and to be compatible with the standards for this project. For each source database (dialysis, transplantation or identification databases), the following must be defined: its exported scheme, its local scheme, the connections between these two schemes, the translator between the local scheme and the exported scheme and vice-versa.

2.3. Interoperability at the network level

The interoperability at the network level requires the integration of computers, operating systems, and software used, while satisfying security requirements and protocol transmission. An innovative solution is characterised by an *n*-tier architecture, integrating a components approach. The first tier is a universal client, of the web-surfer type; the manipulation of the data of the application takes place through the intermediary web pages. The major advantage is in the multi-platform feature and therefore in the portability of the universal client. The second tier is a server made up of two parts, a web interface for communication with the client and a group of business logic components. The web interface is composed of a Java Server Page (JSP), allowing the recovery of html web pages, but also the plug-ins contained within the web pages. Assigned to the interface are a group of software components allowing interaction with the client. The components of the job functions are

software architectures, specialised by subject, which are capable of being easily connected to an integrated system. The architectures are represented under the software components, thereby assuring easy maintenance, but also the capacity for expansion by simple adjunction of components of the same type. These components are described here in the same Java program language, while respecting the overall architecture, using 'Enterprise Java Beans' (EJB). The *n*-tier architecture, by encouraging a separation of the logical, applicative, business, and infrastructure pieces, encourages future development of the system and has the advantages of a tight security and an optimal compatibility.

2.4. Datawarehouse

The datawarehouse model seems well adapted to further improvement of the epidemiological and organisational knowledge of ESRD for public health decision-making. A datawarehouse is actually a collection of data in which the individual is at the heart of the matter. These data are then integrated, made into a historical account, made non-volatile and organised to make up the base of an aid to the decision-making process. Contrary to data from a traditional database, the data from a warehouse are not created 'de novo' by the users, but rather 'calculated' from data of operating bases. In other respects, the warehouse users are essentially analysts and decision makers. In the framework of the REIN program, those responsible for the health decisions made are the users and the doctors of the centres providing the data, who will further be able to visualise the patterns of ESRD on a more global scale at the regional, national, or international levels.

The guaranteed processing by a datawarehouse is regrouped under the name OLAP (On-Line Analytical Processing), which contrasts to the transactional processing of the operating databases called OLTP (On-Line Transactional Processing). The processors are defined by their capacity to strengthen, visualise, and synthesise information in accordance with the analyst's perspective. The information is structured according to multiple axes or dimensions, which makes possible the representation of diverse concepts. This multidimensional structure is most often represented in the form of a hypercube, where each edge represents one dimension from which an activity can be studied.

Starting from a file generated by a medical request folder, we create the components that are capable of the following functionalities: integrating this file into a local reception structure, confirming the coherence of information in accordance with a model, confirming the exhaustiveness of information and the proposal to retrieve additional information, translating the request

into the electronic data exchange format, sending the message to the designator, receiving under the long-distance format, confirming the coherence of information, managing the receipt of the messages and finally, integrating the information into the reception database.

2.5. Interoperability and formats

The exchange of electronic data (EDI) appears to be essential for interoperability. The specification of formatted messages adhering to a syntax and a standardised vocabulary according to international recommendations allow inter-exchangeability but do not exclusively guarantee interoperability. Time and efforts by professionals has made possible an outline of the XML health document formats. Finally, to come close to a standardisation even deeper than the elements of information, a process of standardisation of computer objects is undertaken by the OMG (Object Management Group) with CORBAMED. The modelling under the health-oriented business-objects format represents the next step in regards to the dictionaries; it will allow for the tight integration of applications and information systems. The main steps to supplying the production database begins with the XML messages received, followed by a confirmation of the data given by the care centres or by the web server, and the testing of coherence.

2.6. Interoperability and content

Semantic interoperability requires the management of the data sources' heterogeneity. It stresses the global specification of the semantic, which is driven by the system terms and the representation across a global scheme, and a representative model of the meta-knowledge describing the group of sources. Semantic heterogeneity does not merely concern the data formats of the sources of information. It also involves vocabularies and thesauri, and even bigger, the types of concepts and connections used, as well as the knowledge representation models and the ontologies they implicitly or explicitly refer to.

The standardisation of used terms, vocabularies and concepts and the resort to such reference models as ICD-10, SNOMED or UMLS make up the first step to a solution. Numerous standards and medical references have been developed these past years. An achievement that has already been mastered is represented by the Thesaurus of Nephrology, where the transcribing in ICD-10 is achieved. Ties will be made with the existing work groups (CEN TC251, HL7, EDISANTE).

3. Security, access, transmission and tracking

The combined use of a health professional card and of a secured Internet provided with two levels of encoding, one at the local position and the other one at the transport level. These different services guarantee security for, firstly, a user with protected and/or confidential data, and secondly, for transmissions across the medical computer networks while benefiting from an approved level of security (ITSec E3 fort). The tracking of jobs made on the servers is implemented.

4. Conclusion

The multi-source information system SIMS@REIN is a significant improvement of what already exists and keeps within the general perspective of public health. In the domain of the complex management of ESRD, it is all about taking the best parts from the advancements of

information and communication techniques at the medical level.

It allows the exchange between the acting clinicians, the decision makers and the researchers. The organisation uses a pertinent set of shared medical data. The system is patient-oriented rather than treatment-oriented, as is the current situation. It allows a reconstitution of the steps in healthcare by using a unique ESRD-patient identity. The region is the regrouping unit and is the head of quality control. It favours the cooperation between professionals for the common goals of public health, evaluation and research. On the technological scale, the system is founded on the basis of a datawarehouse and an n -tier architecture that guarantees interoperability. It would allow for more accurate details of the rise of medical expenses in France due to ESRD. It would also provide a precious evaluation tool for the use of resources to be paid for but also to anticipate, at the slightest cost, the increase in the demand of care.

1. Introduction

Le programme *Réseau épidémiologie et information en néphrologie* (REIN) est un programme d'intérêt commun aux malades et aux professionnels de santé et, plus largement, aux acteurs, décideurs et institutions concernés par le traitement de l'insuffisance rénale terminale dans le champ de la santé publique [1]. Ce projet a pour finalité de contribuer à l'élaboration des stratégies sanitaires visant à promouvoir la prévention et les soins développés pour le traitement de l'insuffisance rénale terminale (IRT). Cet objectif stratégique impose une construction favorisant les interactions entre les professionnels, les institutions et les chercheurs (épidémiologistes, économistes de santé), tant au niveau national qu'au niveau régional. Le système d'information multi-sources SIMS@REIN est développé pour réaliser ce programme.

L'absence de système d'information fiable et coordonné limite l'action publique en matière d'IRT. Cela a été souligné de manière récurrente par l'IGAS [2], l'étude IRC 2000 [3], l'Andem [4], la Sanesco [5] et le Haut Comité de santé publique [6], qui s'appuient sur une expertise collective de l'Inserm. Une étude approfondie des systèmes d'information existants, incluant ceux de l'État (SAE, PMSI) et de l'Assurance Maladie (Siam relatif aux prestations et Medicis aux affections de longue durée), réalisée par l'Inserm, à la demande de la Direction générale de la santé, a montré que la France disposait de nombreuses sources d'information,

mais qu'elles étaient parcellaires ou catégorielles et non coordonnées [7].

L'insuffisance rénale terminale (IRT) constitue un enjeu de santé publique, compte tenu de l'évolution des besoins sanitaires de la population et des coûts des traitements. En effet, l'offre de soins est inadaptée, car la demande augmente et se modifie [8]. Les patients sont plus âgés et présentent une comorbidité accrue.

Aucune donnée fiable relative au nombre de patients actuellement pris en charge n'est disponible. Cependant, à partir de données régionales et internationales [9, 10], l'incidence nationale de l'IRT serait voisine 110/Mhab/an et la prévalence voisine de 700/Mhab. Chaque année, plus de 6500 nouveaux patients présentent une insuffisance rénale terminale ; cette incidence augmente de 4 à 5% par an, principalement en raison du vieillissement de la population. Près de deux tiers des 42 000 cas prévalents sont dialysés et un tiers transplantés. La prévalence croît également de 4 à 6% par an. L'évolution des années 90 a été marquée par un vieillissement de la population en IRT, une croissance des comorbidités associées et une mortalité cardiaque et vasculaire accrue [11]. Le coût estimé des traitements par patient est fort élevé. Plus de 10 milliards de francs sont aujourd'hui consacrés au traitement de l'IRT chaque année, soit environ 1,5% du budget de l'Assurance Maladie.

L'objectif du SIMS@REIN est de constituer le support informationnel du REIN. Il vise à coordonner l'inter-opérabilité des acteurs et des systèmes d'infor-

mation avec la mise en place d'un système d'information multi-sources (SIMS). Ce système concerne l'ensemble des patients dialysés et transplantés. Il repose sur un enregistrement continu et exhaustif des nouveaux cas et de leur suivi. Il recueille de manière standardisée un dossier patient minimum, dont les caractéristiques ont été définies avec les professionnels. Il échange des données exhaustives et validées entre les acteurs concernés, dans le respect de la sécurité des données et du droit d'accès des patients, sur le principe d'un Intranet sécurisé.

Ce système gère des flux d'informations multi-sources et son évolution est orientée vers l'interopérabilité entre les trois échelons d'information : local, caractérisé par les centres de traitement (dialyse et transplantation), régional et national. Un échelon supplémentaire est constitué par une coopération renouvelée avec le registre européen.

2. Éléments pour l'interopérabilité

2.1. Le constat

2.1.1. Une informatisation hétérogène

L'informatisation des équipes de néphrologie est hétérogène. Les plus nombreuses n'ont pas encore de gestion automatisée de leurs données. D'autres, en revanche, sont plus avancées et disposent de bases de données informatisées. Elles ont, soit développé des solutions spécifiques, soit adapté des progiciels à des solutions locales. Entre ces deux catégories, des équipes ont acquis un ou plusieurs logiciels pour la gestion de leur système d'information, et sont dépendantes de leur fournisseur pour toute évolution ultérieure. La limitation des moyens locaux est souvent conjuguée à une inquiétude des responsables des centres quant à une éventuelle perte d'autonomie dans le cadre d'une approche favorisant l'interopérabilité, donc l'accès à leurs données. Ces éléments doivent être pris en compte pour l'élaboration de propositions évolutives.

Construire un nouveau système d'information qui remplacerait les systèmes existants est en général hors de la portée des centres, tant du point de vue humain que financier. Techniquement, deux solutions sont possibles : soit réaliser des programmes ad hoc au fur et à mesure des besoins, soit mettre en place un système d'information multi-sources (SIMS), qui conserve les applications et les données existantes, mais étend leurs possibilités. La première solution serait fonctionnelle à court terme ; cependant, à moyen ou à long terme, l'évolutivité et la cohérence globale du système deviendraient rapidement difficiles à maîtriser. La deuxième solution est initialement plus coûteuse, car elle mobilise

plus de ressources humaines et informatiques, mais elle permet une meilleure maîtrise de l'évolution du système et de la consolidation des données.

2.1.2. Des sources d'information diverses

Un système d'informations multi-sources (SIMS) s'appuie sur un ensemble de sources (de données et de connaissances) distribuées et hétérogènes (diversité de technologie, d'accès, de perception et de représentation), qui proposent un sous-ensemble de leurs données (schéma exporté) au partage, éventuellement conditionné par des droits. S'il doit être à même de répondre à un ensemble de requêtes spécifiques aux problèmes qui sont à son origine, il doit aussi demeurer ouvert, pour permettre l'intégration de nouveaux composants ou répondre à de nouveaux besoins. Les entrepôts de données (*datawarehouse*), les bases interopérables (intégrées ou fédérées), les applications web sur Intranet ou Internet sont des exemples de SIMS [12]. Lorsque les utilisateurs disposent d'un SIMS, ils n'ont qu'un seul système à apprendre (fonctionnalités, interfaces, termes, etc.), la navigation entre bases est transparente, et l'information disponible est plus riche et plus sûre.

2.1.3. L'hétérogénéité des bases de données

Une des grandes difficultés de la mise en œuvre de SIMS provient de l'hétérogénéité des bases composantes [13, 14]. L'hétérogénéité se manifeste à différents niveaux. On distingue classiquement le niveau *système*, caractérisé par des différences de systèmes d'exploitation (Unix, NT...), de protocoles (TCP/IP, IPX...) ou de systèmes de gestion de bases de données (Oracle, Access, Omnis, fichiers...), et le niveau *sémantique*, où l'on fait référence aux différentes manières dont les entités du monde réel sont modélisées et représentées. L'hétérogénéité des systèmes est gérée par les outils de *middleware* et par des extensions apportées à leur système par les concepteurs de SGBD. L'hétérogénéité sémantique exprime les différences dans la signification et l'utilisation des données, qui rendent difficile l'identification des relations qui existent entre les objets d'une même base ou de bases interconnectées. Les schémas conceptuels sont différents, un même nom peut être utilisé pour dénoter des concepts (attributs, classes, relations) différents ; le même concept peut être désigné par des noms différents, le type et le format de représentation de données semblables peuvent être différents. Il s'agit donc d'intégrer des objets sémantiquement semblables, mais structurellement différents.

L'interopérabilité des systèmes d'information peut être définie comme la capacité de plusieurs systèmes d'information, non seulement à échanger de l'information (interopérabilité fonctionnelle), mais aussi à utiliser l'information échangée (interopérabilité sémantique).

2.2. Interopérabilité au niveau sémantique

On peut considérer quatre grandes familles d’architectures logicielles pour la résolution de l’hétérogénéité sémantique [15].

2.2.1. Approche traducteur

Pour une tâche définie, par exemple dans le cadre d’une étude épidémiologique, et à chaque couple de bases qui doivent échanger des données, on associe un traducteur (Fig. 1). Dans le pire des cas, pour n bases de données, on aura besoin d’écrire environ n^2 traducteurs par tâche. Cette architecture s’applique bien aux domaines où les tâches et les flux de données sont connus et en nombre limité, ce qui est par exemple le cas des applications d’alimentation d’un entrepôt de données (*datawarehouse*). Cette architecture pose problème dans des domaines ouverts, car la prise en compte de toute nouvelle tâche nécessite la réécriture de nouveaux traducteurs. Dès lors, le maintien des divers traducteurs peut devenir une tâche à part entière.

2.2.2. Approche intégrée

Les méta-connaissances des source locales et l’ontologie du domaine sont regroupées dans un serveur central. Toute requête au SIMS, requête exprimée dans le langage du SIMS, est analysée par ce serveur, puis traduite et envoyée aux sources concernées. Les résultats

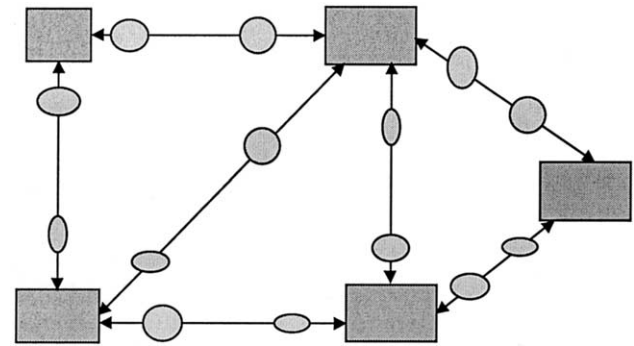


Fig. 1. Approche traducteur. Les rectangles représentent les bases sources et les ovales des traducteurs.

fournis par ces sources sont alors retraduits dans le serveur central, fusionnés et renvoyés à l’utilisateur (Fig. 2). La centralisation des connaissances du SIMS présente deux principaux avantages : la sémantique du SIMS étant centralisée, elle est plus facilement maintenue et la concurrence d’accès est plus facilement gérable. En contrepartie, les traducteurs vers les divers systèmes étant aussi centralisés, les temps de réponse peuvent poser problème.

2.2.3. Approche complètement décentralisée

Dans cette approche, qui est à l’opposé de l’approche précédente, chaque base locale contient un compilateur du langage du SIMS et un module de communication.

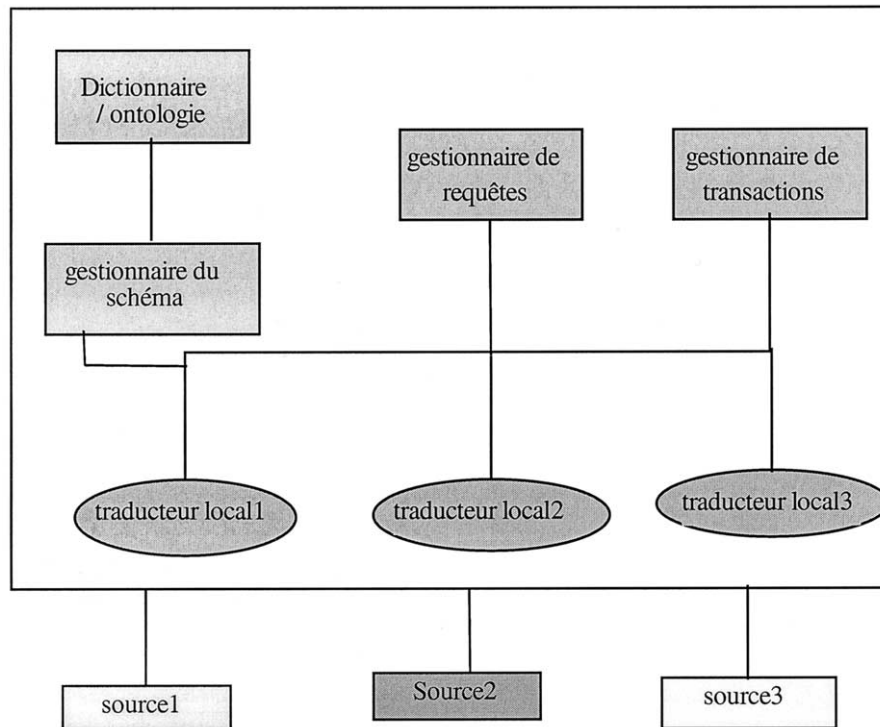


Fig. 2. Approche intégrée. Les rectangles représentent les bases sources et les ovales des traducteurs.

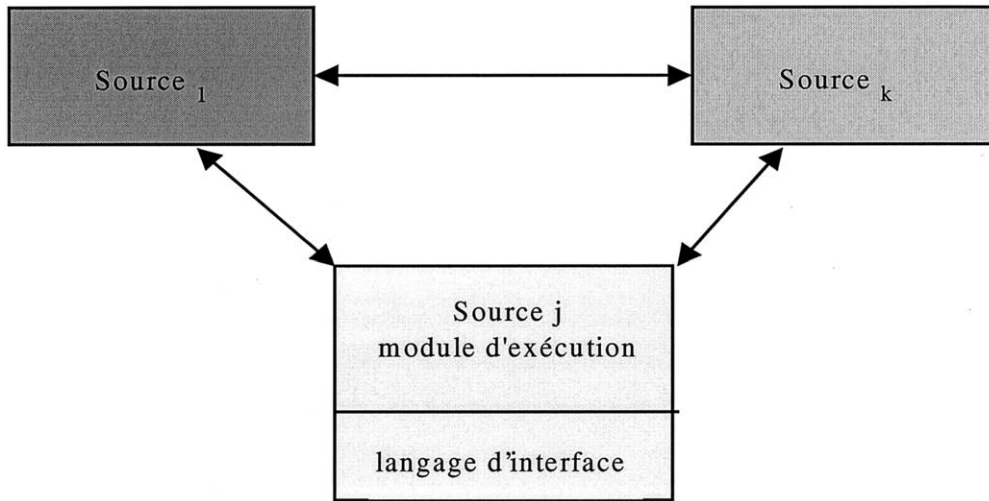


Fig. 3. Approche décentralisée.

La résolution des conflits lors de l'exécution d'une requête est réalisée par des applications spécifiques (Fig. 3). Pour favoriser la coopération inter-bases, on conseille fortement que l'intégration de bases locales se fasse en utilisant une ontologie de domaine. Cette solution permet une plus grande flexibilité dans la définition des méta-connaissances locales. En revanche, elle est à l'origine de modules d'exécution locaux plus complexes. De plus, la gestion de la concurrence est plus difficile que dans le cas centralisé. Enfin, il n'y a pas de contrôle de la résolution de l'hétérogénéité sémantique, car on ne fait que « conseiller » d'utiliser l'ontologie de domaine.

2.2.4. Approche middleware

Un *middleware* dispose d'un certain nombre de services, qui utilisent en particulier un détecteur de conflits, l'ontologie de domaine et les bibliothèques de

description des méta-connaissances locales, aussi dénommées contextes locaux. Ainsi, lorsqu'un client envoie une requête (dans son propre langage et non plus dans le langage du SIMS), le *middleware* résout les conflits, convertit la requête en requête(s) locale(s) et les envoie aux bases sources. Celles-ci exécutent la requête et envoient le résultat au *middleware* qui, après conversion, les renvoie au client (Fig. 4). Un avantage de cette solution est que la résolution de l'hétérogénéité sémantique a été rendue transparente pour les utilisateurs, car ceux-ci continuent d'utiliser leur propre langage. La contrepartie est que cela exige la mise en place des services pour l'exploitation des connaissances (les services du *middleware*).

En résumé, l'approche *traduction* convient bien à l'alimentation de l'entrepôt de données, qui constitue les bases régionales et nationale du projet. Elle peut être

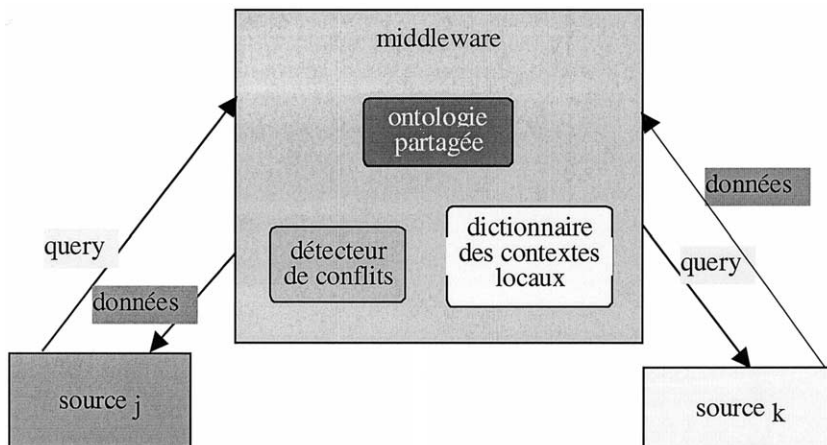


Fig. 4. Approche middleware.

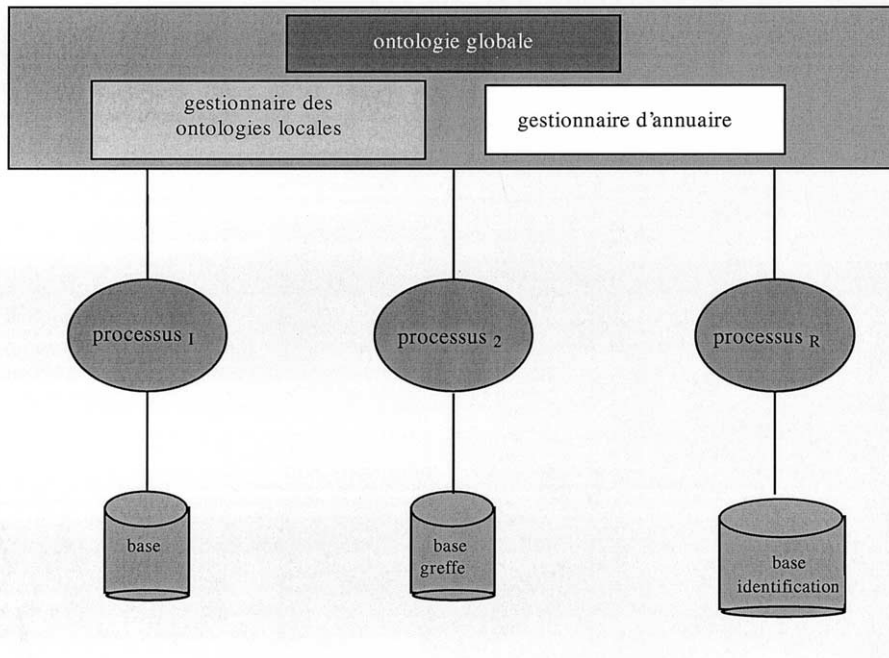


Fig. 5. Architecture pour l'interopérabilité.

déployée directement. En revanche, elle présente des limites dans ses capacités d'évolution.

2.2.5. Mise en œuvre

Toute base peut fonctionner selon deux modes distincts : un fonctionnement local et un fonctionnement en coopération, dans lequel l'utilisateur envoie (directement ou indirectement) une requête. Celle-ci est traduite en utilisant le traducteur *local/exporté* et envoyée au *middleware*, qui localise les serveurs cibles, décompose le message et renvoie à chaque serveur potentiel le message adéquat. Chaque serveur utilise alors le traducteur *exporté/local* pour réécrire le message en termes du schéma local et exécute la requête ainsi générée. Les données suivent alors un cheminement inverse pour être enfin présentées à l'utilisateur. À titre d'exemple, nous proposons l'architecture présentée sur la Fig. 5.

Par rapport à la solution présentée sur la Fig. 4, une partie de l'information sur les bases sources (dans une première étape, les bases sources considérées sont la base d'événements en dialyse, la base d'événements de la greffe et la base d'identification) est décentralisée, ce qui rend l'intégration ultérieure d'une nouvelle base source plus flexible. En privilégiant une approche déclarative de la méta-connaissance des diverses bases, on favorise l'intégration des messages nécessaires à de nouvelles applications.

Un processus de communication entre une base et le *middleware* est essentiellement composé d'un

gestionnaire de correspondances entre les termes du schéma local et ceux du schéma exporté et de deux traducteurs qui, en fonction du gestionnaire de correspondances, assurent la traduction entre le schéma exporté et le schéma local, et vice-versa (Fig. 6).

Cette solution suppose aussi une décentralisation de la compétence. Les administrateurs locaux sont fortement impliqués dans la mise en place des processus de communication avec le *middleware*, ce qui est un gage de réussite du projet. Enfin, elle permet également d'envisager que l'utilisateur ne saisisse qu'une seule fois ses données, et que celles-ci soient transmises à sa base personnelle et à d'autres bases.

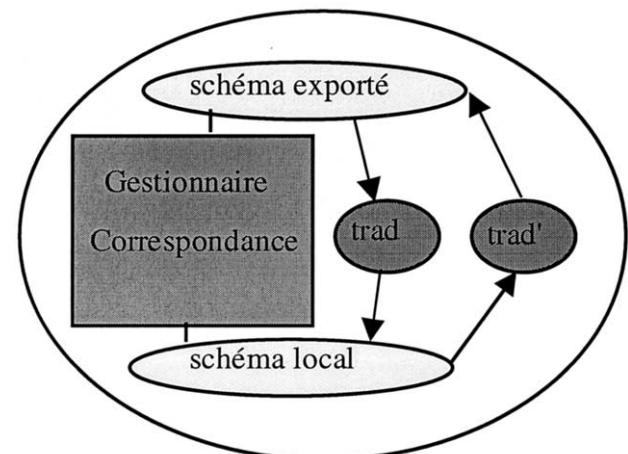


Fig. 6. Composants d'un processus de communication.

La mise en place de cette solution nécessite deux types de définitions. Il faut définir l'ontologie globale, aussi dénommée schéma global ou schéma canonique. Elle s'appuiera sur le noyau préalablement défini [16] et devra aussi disposer des informations nécessaires à des échanges entre bases locales, ainsi qu'être compatible avec les standards retenus pour ce projet. Il faut définir aussi, pour chaque base source (base d'événements dialyse, base d'événements greffe, base d'identification), son schéma exporté, son schéma local, les correspondances entre ces deux schémas, les traducteurs entre le schéma local et le schéma exporté et vice-versa.

2.3. Interopérabilité au niveau du réseau

Différentes approches permettent d'améliorer le transfert des informations, depuis les niveaux les plus bas, établissant des liens forts entre les bases de données, jusqu'aux liens les plus lâches. Elles confèrent au réseau une certaine « intelligence ». La partie intermédiaire, dite de *middleware*, peut se matérialiser schématiquement sous trois formes diverses. Il peut s'agir :
 – de liaisons fortes avec les serveurs SQL (langage de requêtes structurées), qui permettent la mise en œuvre d'échanges de données par procédures stockées [17, 18] ; elles répondent à critères stricts (propriétés ACID par exemple, pour atomicité, cohérence, isolation et durabilité) et imposent un comportement synchrone entre les clients et le (ou les) serveur(s) distant(s) ;
 – de liaisons dont les liens sont faiblement couplés, à la manière d'un courrier, par opposition à un appel téléphonique ; la transcription informatique de ce comportement est appelé un MOM (pour *Message-Oriented Middleware*) [19] ; le mode de transfert est asynchrone, c'est-à-dire que le récipiendaire peut être absent au moment de l'envoi ; il recevra le message dès qu'il sera

à nouveau présent ; cette caractéristique différencie ce mode de transfert des approches synchrones, qui nécessitent la présence du récipiendaire lors de l'envoi du message ;

– enfin, d'autres technologies à objets distribués [20–22], qui proposent une possibilité d'échanges en incluant des services annexes ; outre la sécurité ou les comportements, elles permettent d'inclure des comportements exécutoires aux données transmises grâce à des objets réseau (*network object aware*) ; ceux-ci sont transférés et exécutés ; c'est, par exemple, une mise en forme graphique.

2.3.1. Architecture de type n-tiers

L'interopérabilité au niveau du réseau nécessite l'intégration des machines, des systèmes d'exploitation et des logiciels mis en œuvre, tout en satisfaisant aux impératifs de sécurité et de transmission de protocole. Une solution émergente est caractérisée par une architecture de type *n*-tiers, intégrant une approche par composants (Fig. 7). Le premier tiers est un client universel, de type navigateur, ce qui implique que la manipulation des données de l'application s'effectue par l'intermédiaire de pages web. L'avantage majeur de cette solution se trouve dans le caractère multiplateformes et donc dans la portabilité du client universel. Le deuxième tiers est un serveur composé de deux parties, une interface web de communication avec le client et un ensemble de composants dans la partie fonctions métiers (*business logic*). L'interface web se compose d'un Java Server Page (JSP), permettant la récupération dynamique des pages html, mais aussi des extensions (*plug-ins*) contenues dans les pages. Il lui est associé un ensemble de *servlets* (composants logiciels) permettant l'interaction avec le client. Les composants de la partie fonctions métiers sont des architectures logicielles spécialisées par sujet, qui peuvent être faci-

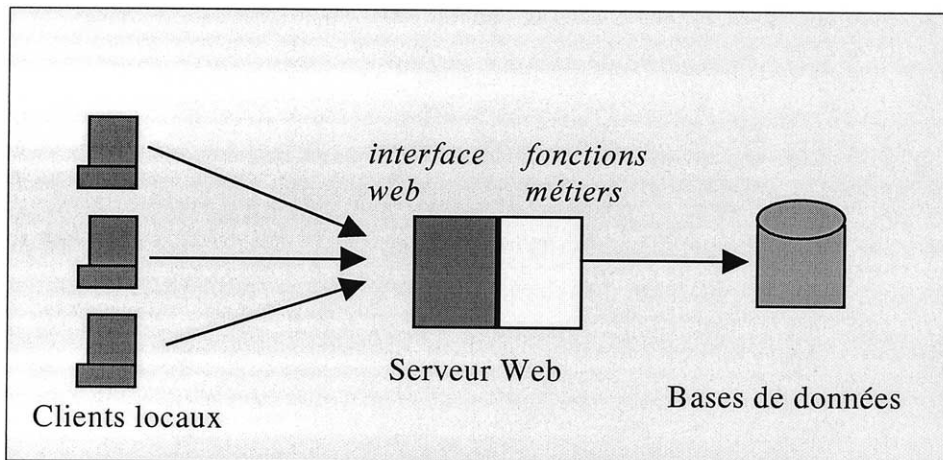


Fig. 7. Architecture de type 3-tiers.

lement connectées à un système intégré. Ils se présentent sous la forme de composants logiciels, assurant ainsi une maintenance aisée, mais aussi une capacité d'extension par simple adjonction de composants de même type. Ces composants sont décrits ici dans le même langage de programmation Java, respectant ainsi l'unité générale de l'architecture, à savoir les *Entreprise Java Beans* (EJB). Ces composants prennent en charge les transactions vers le troisième tiers, à savoir les bases de données.

L'architecture *n*-tiers, en préconisant une bonne séparation des partitions logique, applicative, métier et infrastructure, favorise l'évolution ultérieure du système et présente l'avantage d'une sécurité et d'une compatibilité optimales.

2.3.2. Mise en œuvre

À titre d'exemple, dans notre programme de développement de SIMS@REIN, nous avons retenu une architecture de type *n*-tiers isolant le suivi des données de leur exploitation.

Le navigateur intègre des pages de type hypertexte (html, pour *hypertext markup language*), présentant aux utilisateurs finaux des formulaires, des requêteurs ou même des interfaces de présentation dans le cadre de retour d'information pour le suivi épidémiologique. On peut associer aisément à ces navigateurs des cartes, telle que la Carte des professionnels de santé, qui assure l'identification de l'utilisateur ainsi que le cryptage/décryptage des données transmises/reçues.

L'interface web encapsule les pages html et récupère les données générées par le client. Sur son versant « troisième tiers », on trouve les fonctions métiers, qui procurent les interfaces nécessaires pour les fonctionnalités d'alimentation, d'identification, d'intégration et du suivi événementiel des informations sur les patients.

Au niveau national, l'identification du patient dans SIMS-REIN est unique via le serveur d'identité. Ce serveur a comme objectif de positionner les informations définissant une personne physique, d'attribuer un identifiant permanent de l'identité, de fournir une gestion d'identifiant temporaire de l'identité, voire de stocker des identifiants permanents émanant d'applications externes. Il permet aussi de qualifier de façon unique les intervenants de santé, puisqu'ils sont considérés comme des personnes physiques.

Le serveur d'événements possède la capacité d'intégrer des ensembles d'informations venant des centres de soins et d'enrichir l'histoire sanitaire et clinique du patient. Ceci permet d'avoir une vision globale des événements sur l'axe temporel. Un événement peut être vu comme une synthèse d'informations médicales ou autres, sous la responsabilité d'un auteur. Il peut aussi

illustrer le déplacement géographique du patient dans le temps. L'événement possède un état défini à un instant donné et se compose de trois parties : le contenant (définition du type d'événement, date de référence, libellé, auteur, type d'objet décliné), le contenu d'informations synthétiques (résumé) et l'objet actif (contenu d'information complet et les méthodes associées). Une analyse des différents événements existants dans le système d'information est mise en œuvre. Une analyse du contenu des états ainsi que celle des données et des méthodes afférentes à l'objet manipulé est également développée.

Enfin, le troisième tiers se compose d'une base d'identification, d'une base de production et d'un entrepôt de données.

2.4. Entrepôt de données

2.4.1. Modèle

Le modèle des entrepôts de données s'avère bien adapté pour répondre à l'objectif majeur du projet, qui est l'amélioration des connaissances épidémiologiques et organisationnelles pour la prise de décision en santé publique. En effet, un entrepôt de données est une collection de données, dont l'individu est le cœur ; celles-ci sont intégrées, historisées, non volatiles, et organisées pour constituer le support d'un processus d'aide à la décision [23]. Les entrepôts de données se distinguent des bases de données traditionnelles par la manière dont leurs données sont produites et utilisées. Contrairement aux données d'une base traditionnelle, les données d'un entrepôt ne sont pas créées « de novo » par les utilisateurs, mais « calculées » à partir des données des bases opérationnelles. Par ailleurs, les utilisateurs d'un entrepôt sont essentiellement des analystes et des décideurs ; ils n'appartiennent donc pas aux mêmes catégories que les utilisateurs des bases opérationnelles. Dans le cadre du projet REIN, les utilisateurs sont les responsables des décisions de santé, et aussi les médecins des centres fournisseurs de données, qui pourront ainsi se positionner par rapport à une situation plus globale : régionale, nationale, voire internationale.

2.4.2. OLAP

Les traitements assurés par un entrepôt de données sont regroupés sous le terme OLAP (*On-Line Analytical Processing*) [24], qui s'oppose au traitement transactionnel des bases de données opérationnelles appelé OLTP (*On-Line Transactional Processing*). Ces traitements sont caractérisés par leur aptitude à consolider, visualiser et synthétiser des informations reflétant la vision de l'analyste. L'information est structurée selon

plusieurs axes ou dimensions, ce qui permet de représenter des notions diverses, comme le temps ou la localisation géographique. Une donnée est un point dans un espace à plusieurs dimensions. Cette structure multidimensionnelle est le plus souvent représentée sous la forme d'un hypercube, dont chaque arête représente une des dimensions selon lesquelles on veut étudier l'activité. Un hypercube est composé de cellules contenant une ou plusieurs mesures. Une mesure est une donnée quantitative relative à l'activité étudiée. Le nombre de cas incidents ayant un âge compris entre 40 et 45 ans et habitant une ville donnée est un exemple de mesure. En pratique, un hypercube est toujours présenté au moyen de projections sur au plus trois dimensions, donc par un cube.

Les dimensions permettent aussi d'exprimer différents points de vue sur les données. Par exemple, selon les besoins, la dimension temporelle pourra être considérée à l'échelle du jour, du mois, du trimestre, de l'année ou de la décennie. De même, l'axe géographique permet d'exprimer différents niveaux de regroupement : par ville, par région, par zone urbaine ou rurale, etc. Outre ces deux axes, qui sont classiques dans les entrepôts de données, on sera amené à considérer les tranches d'âge, les catégories socio-professionnelles, les comorbidités associées, la mortalité.

2.4.3. Mise en œuvre

À partir d'un fichier généré par un dossier de recueil médical, on crée les composants capables de remplir les fonctionnalités suivantes : intégrer ce fichier dans une structure d'accueil locale, valider la cohérence des informations par rapport à un modèle, valider l'exhaustivité des informations et proposer la saisie complémentaire d'informations, traduire le recueil au format d'échange des données électroniques (voir plus bas), envoyer le message au destinataire (sécurisation), le réceptionner dans la structure d'accueil distante, valider la cohérence de l'information et gérer les acquittements des messages, enfin intégrer les informations dans la base de réception.

Dans SIMS@REIN, l'alimentation du *datamart* se fait périodiquement à partir des données consolidées de la base d'événements. Les composants sont capables de remplir les fonctionnalités suivantes : accéder à la base d'événements de dialyse, à la base d'événements de greffe et au serveur d'identité pour récupérer les (nouvelles) données consolidées ; dériver les données nécessaires à l'alimentation du *datamart* ; agréger et intégrer ces données dans un nouvel état de l'entrepôt de données ou dans une nouvelle version d'un état existant, calculer les extensions des vues persistantes, vérifier la cohérence de l'information intégrée, gérer les

acquittements des messages, intégrer les messages dans le dictionnaire de l'entrepôt de données.

2.5. Interopérabilité au niveau des formats

2.5.1. Formats de données

En France, la normalisation des données et des systèmes informatiques est coordonnée par une commission de normalisation de l'informatique de santé à l'Afnor. Elle traite en particulier de formats de données de différents types, du dossier médical, de la sécurité. Cette commission est le miroir français du comité technique spécialisé santé au niveau européen (Technical Committee 251), dont les groupes de travail et les équipes de projet élaborent les normes à partir d'une modélisation d'ensemble du système socio-sanitaire [25, 26].

L'échange de données électroniques (*Electronical Data Interchange* ou EDI) apparaît comme une condition nécessaire de l'interopérabilité. La spécification de messages structurés respectant une syntaxe et un vocabulaire normalisés selon des recommandations internationales (UN/EDIFACT, ASTM) permet l'interéchangeabilité, mais elle ne garantit pas, sans l'exclure, l'interopérabilité.

Des travaux ont permis d'esquisser des structures de documents XML pour la santé (groupe de travail du CEN, Édisanté). Enfin, pour approcher une normalisation encore plus profonde des éléments d'information, une démarche de standardisation des objets informatiques est entreprise par l'OMG (*Object Management Group*) avec Corbamed [27, 28]. La modélisation sous forme d'objets-métiers, spécialisés santé, représente l'étape suivante par rapport aux dictionnaires, et elle permettra d'intégrer de façon étroite des applications et des systèmes d'information.

2.5.2. Mise en œuvre

Afin de transmettre les informations des centres de soins contenus dans leur base de données locale vers le serveur de données de production, nous nous appuyons sur des structures définies d'échanges de données. L'analyse des différents types de messages et leur validation est réalisée avec tous les partenaires. Dans le cas où le recensement des informations ne peut s'appuyer sur des types de messages existants, on écrit la structure et la sémantique des messages sous la norme XML (langage de méta-description) en définissant la DTD (*Document Type Definition*) et les schémas de ces échanges. Ces spécifications seront rendus publiques. Dès l'instant où un standard d'échange existe, comme par exemple HL7 (au format XML) pour l'identification des patients, il sera utilisé d'office. Ces

messages pourront aussi servir de support d'échange entre le serveur web et les serveurs applicatifs.

Dans le cadre de l'alimentation de la base de données de production à partir des messages XML reçus, on réalise une validation des données émises par les centres de soins ou par le serveur web, et on en teste la cohérence. On teste également l'interprétation des données XML et leur intégration dans le système d'information grâce aux fonctions métiers du serveur d'identité et d'événements. Ces traitements s'effectuent au niveau du système serveur applicatif (2-tiers). Enfin, une gestion d'acquiescement est mise en œuvre.

2.6. Interopérabilité au niveau du contenu

2.6.1. Hétérogénéité sémantique

L'interopérabilité sémantique nécessite de gérer l'hétérogénéité des sources de données. Elle exige la spécification globale de la sémantique véhiculée par les termes du système et sa représentation à travers un schéma global (extensible) et un modèle de représentation des méta-connaissances décrivant l'ensemble des sources. Ainsi, la description hiérarchique des messages dans le projet *Health Level Seven* (HL7) intègre les règles de construction syntaxique des messages dans le modèle conceptuel plus général de l'interopérabilité [29]. L'hétérogénéité sémantique ne concerne pas seulement les structures de données (nom de champs, attributs, relations) des sources d'information. Les sources d'informations ne décrivent pas forcément le même domaine et peuvent l'aborder avec une approche ou des finalités d'exploitation différentes. L'hétérogénéité sémantique concerne donc aussi les vocabulaires et les thesauri employés et, plus largement, les types de concepts et de relations utilisés, ainsi que les modèles de représentation des connaissances et les ontologies auxquelles ils se réfèrent, implicitement ou explicitement. Le problème de l'hétérogénéité sémantique a été identifié et des solutions à type de médiateur sémantique proposées [30].

2.6.2. Standardisation de la terminologie

La standardisation des termes, des vocabulaires et des concepts utilisés ainsi que le recours à des modèles de référence tels la CIM-10, Snomed ou UMLS [31] constituent un premier niveau de solution. Dans le domaine de la dialyse et de la transplantation, le thesaurus de néphrologie [32] constitue une base importante de vocabulaire, agréée par le pôle d'expertise et de référence national des nomenclatures de santé (PERNNS).

Toutefois, des solutions à type de standardisation ne constituent pas une solution suffisante si l'on tient compte des sources de données telles qu'elles existent,

à moins de faire peser des contraintes lourdes sur les administrateurs de ces bases de données. Ce sont alors des solutions de type transcodage ou traduction qu'il faut envisager [33].

2.6.3. Vocabulaires et terminologies

De nombreux standards et référentiels médicaux ont été développés au cours de ces dernières années [34–36], mais beaucoup reste à faire [37]. Une première étape a été de recenser l'existant et d'évaluer son adaptation au domaine de l'insuffisance rénale. Parmi les standards utilisés dans d'autres contextes (imagerie, biologie), certains sont a priori réutilisables [38]. D'autres sont à créer. Un acquis déjà maîtrisé est représenté par le thesaurus de néphrologie, qui constitue une base importante de termes du domaine, dont le transcodage en CIM-10 est acquis [39]. Des liens seront tissés avec les groupes de travail existants (CEN TC251, HL7, Édisanté).

La définition a priori de référentiels peut constituer une tâche très longue, même pour un micro-domaine, alors que les sources d'information n'utilisent le plus souvent qu'un champ restreint du « vocabulaire » offert par les thesauri. De plus, les gisements de données médicales comportent souvent un volume non négligeable de données manquantes. Une approche pragmatique consiste donc à constituer un corpus d'enregistrements de patients pour guider les choix en terme d'étendue et de contenu des termes (concepts) utiles. Ce vocabulaire « minimum » sera au besoin complété pour disposer des informations requises au niveau de la base d'exploitation.

2.6.4. Formalismes de représentation des connaissances et graphes conceptuels

Le modèle des graphes conceptuels est un formalisme de représentation des connaissances [40–42]. Il a fait l'objet de différents travaux dans le domaine de l'information médicale et de l'échange de données électroniques [43–47]. Ce modèle présente l'avantage de fournir un ensemble d'opérations formelles sur les graphes : copie (graphe identique au graphe de départ), restriction (remplacement d'un concept par une de ses spécialisations), généralisation (remplacement d'un concept par un sur-type plus générique), jonction (de deux graphes sur un concept commun) et simplification (élimination d'une arête dupliquée).

Les choix réalisés en terme de modèle de représentation des connaissances et de méta-modèle de représentation des structures de données interagissent avec le contenu, voire le format, des messages échangés. Par ailleurs, la complexité des solutions à mettre en œuvre dépend aussi du degré requis d'autonomie des sources d'information ainsi que de l'ouverture et de l'évolutivité du système d'information multi-sources.

3. Sécurité

3.1. Sécurisation des accès et des transmissions

L'emploi conjoint d'une carte de professionnel de santé (CPS) et d'un Intranet sécurisé fournit deux niveaux de cryptage, l'un au niveau du poste local et l'autre au niveau du transport. Ces différents services assurent la sécurisation, d'une part, des accès d'un utilisateur à des données protégées et/ou confidentielles et, d'autre part, des transmissions à travers les réseaux informatiques du système de santé en bénéficiant d'un niveau de sécurité homologué (ITSec E3 fort).

Pour assurer la recherche de son correspondant, un annuaire est indispensable dans les échanges électroniques de ses clefs publiques (pour vérifier sa signature ou lui envoyer des données chiffrées), ainsi que pour consulter la liste d'opposition. Le chiffrement peut-être effectué directement sur le poste de travail (données, messages...), avant de l'envoyer en pièce jointe par toute messagerie du commerce. Seul le destinataire peut le déchiffrer sur son poste. Ce chiffrement de bout en bout, indépendant des environnements et des réseaux utilisés, présente l'avantage d'une sécurité et d'une compatibilité optimales.

3.2. Traçabilité

La traçabilité des actions sur les serveurs est indispensable. Toute manipulation des données réalisée au travers des serveurs d'identité et d'événements sur la base de production pourra être identifiée en fournissant le type d'action réalisé (création, modification, suppression, visualisation...), l'utilisateur manipulateur, le type d'information consulté (patient, compte rendu...), l'identification de l'objet activé et la date de réalisation.

4. Conclusion

Tels sont, schématiquement présentés, les principaux éléments de l'évolution des systèmes d'information conditionnant le choix des solutions techniques appropriées. Parmi les développements envisagés pour notre SIMS, certains seront opérationnels à court terme, allant jusqu'au prototype. D'autres se limitent à des choix de modèle, de formalisme de représentation ou à des spécifications, de telle sorte qu'au pragmatisme des solutions opérationnelles proposées s'associe une identification des limites consenties, sans obérer l'avenir.

Le système d'information multi-sources SIMS@REIN constitue une amélioration significative de l'existant et s'inscrit dans une perspective de santé publique. Dans le domaine complexe de la gestion de l'IRT, il s'agit de tirer le meilleur parti de l'avancement

des techniques de l'information et de la communication au service de la santé.

Il permet l'échange entre les acteurs cliniciens (dialyse et transplantation), les décideurs (ministère des Affaires sociales, Cnam, ARH, EfG), et les chercheurs (universités, Inserm). L'organisation repose sur un recueil minimum partagé d'informations. Le système est centré sur le patient (et non plus sur les traitements comme actuellement). Il permet de reconstituer la filière de soins par l'utilisation d'un identifiant patient unique des insuffisants rénaux terminaux. Il est articulé par région, unité de regroupement et siège du contrôle de qualité. Il favorise la coopération entre professionnels pour des objectifs complémentaires de santé publique, d'évaluation et de recherche. Sur le plan technologique, le système est fondé sur un entrepôt de données et une architecture de type *n*-tiers, qui assurent l'interopérabilité.

La mise en place d'un tel dispositif est utile si l'on considère que l'offre de soins pour l'IRT ne peut plus être gérée « en boîte noire ». Il permettrait de préciser en France l'amplitude de la progression de la prise en charge de l'IRT. Il fournirait un outil précieux d'évaluation de l'utilisation des ressources pour faire prendre en charge, mais aussi anticiper, au moindre coût, l'augmentation de la demande de soins. Dans le cas particulier de l'IRT, du fait que l'accès aux soins n'est pas limité a priori, l'étude des patients traités présente non seulement un intérêt du point de vue de la santé publique, mais également du point de vue de la surveillance des risques infectieux et iatrogènes, ainsi que de la recherche, notamment à visée de prévention et d'économie de santé. Il s'agit de mettre en place un système d'information au service de la santé publique.

Les membres du projet participent ainsi à la normalisation des échanges d'informations dans le domaine de la santé. Les spécifications qui seront produites seront publiques et diffusées auprès des acteurs de la dialyse et de la transplantation et auprès des éditeurs de logiciels.

Enfin, il est important de souligner que plus de 60 pays, dont les États-Unis, le Canada, l'Australie, le Japon, l'Allemagne, la Grande-Bretagne, la Hollande ou la Belgique disposent depuis plusieurs années d'un enregistrement national exhaustif des patients en IRT, mis en place à la demande des pouvoirs publics, en collaboration avec les néphrologues, pour le plus grand bénéfice des différents partenaires [1]. Notre retard en la matière pourrait être comblé par une approche technique plus avancée que les autres pays et fondée sur une collaboration raisonnée des principaux acteurs.

Remerciements. Ce programme a bénéficié en 1999 d'un financement du ministère de l'Éducation nationale, de la Recherche et de la Technologie, dans le cadre du programme Télémedecine N°99B 0625. Il est réalisé en coopération avec l'Établissement français des greffes, l'Inserm U170, l'entreprise R2i Santé, Cegetel.rss, le GIP CPS et l'association Édisanté. Mlle Astrid de Bantel est vivement remerciée pour l'aide qu'elle a apporté à la rédaction de la version anglaise du manuscrit.

Références

- [1] P. Landais, B. Stengel, C. Fumeron, C. Jacquelinet, L'insuffisance rénale terminale traitée en France – Épidémiologie et système d'information, *Méd. Théor.* 4 (1998) 533–542.
- [2] Inspection générale des affaires sociales, Enquête sur la dialyse rénale, rapport n° 94.092, 1994.
- [3] Société de néphrologie–Sanesco, IRC : le livre blanc de la néphrologie, *Néphrologie* 17 (1996) (2000) 1–103.
- [4] Agence nationale pour le développement de l'évaluation médicale, Indications de l'épuration extra-rénale dans l'insuffisance rénale chronique terminale, rapport Andem, Paris, 1996, pp. 224.
- [5] R. Cash, P. Fournier, C. Meyer, Schéma de régulation pour la prise en charge de l'insuffisance rénale chronique au stade terminal., Sanesco, Paris, 1996, pp. 150.
- [6] Haut Comité de la santé publique, Avis sur la prévention de l'insuffisance rénale chronique et son diagnostic précoce, Éditions ENSP, collection Avis et Rapports, Paris, 1997, pp. 27.
- [7] B. Stengel, P. Landais, Recueil d'informations sur la prise en charge de l'insuffisance rénale terminale, étude de faisabilité, Direction générale de la santé, Inserm, Paris, Convention d'étude n° 17 96 012, 1997, pp. 120.
- [8] P. Landais, L'insuffisance rénale terminale, un enjeu de santé publique, *Méd. Théor.* 4 (1998) 531–532.
- [9] P. Jungers, G. Choukroun, C. Robino, P. Taupin, M. Labrunie, N.K. Man, P. Landais, Épidémiologie de l'insuffisance rénale terminale en Île-de-France : enquête coopérative prospective 1998, *Nephrol.* 5 (2000) (1998) 239–246.
- [10] P. Landais, Trends in the epidemiology of end-stage renal disease over the last decade, in: P. Cochat (Ed.), *Transplantation and Changing Management of Organ Failure*, Kluwer Academic Publishers, Dordrecht, 1999, pp. 3–28.
- [11] SIMS@REIN. Un système d'information multi-sources pour l'insuffisance rénale terminale, ministère de la Recherche et de la Technologie, Réseau national Technologie pour la santé, Appel d'offres RNTS, Paris, 2000.
- [12] W.H. Inmon, R.D. Hackathorn, *Using the Data Warehouse*, Wiley–QED publication, Londres, 1994, pp. 512.
- [13] C. Roncancio, Interopérabilité entre SGBD : systèmes fédérés et systèmes multibases, *Techn. Sci. Informat.* 13 (1994) 385–419.
- [14] A. Sheth, J. Larson, Federated systems for managing distributed, heterogeneous, and autonomous databases, *ACM Comput. Surv.* 22 (1990) 183–235.
- [15] A. Elmagarmid, M. Rusinkiewicz, A. Sheth, Heterogenous and autonomous database systems, Morgan Kaufmann, Londres, p. 1999.
- [16] P. Landais, L'insuffisance rénale terminale en France. Offre de soins et prévention, *Presse Méd.* 31 (2002) 176–185.
- [17] G. Gardarin, O. Gardarin, *Le client-serveur*, Eyrolles, Paris, 1996.
- [18] G. Coulouris, J. Dollimore, T. Kindberg, *Distributed Systems: Concepts and Design*, 2nd edition, Addison-Wesley, New York, 1994, pp. 672.
- [19] M. Shoffner, Write your own general-purpose, message oriented middleware, *JavaWorld* (www.javaworld.com/javaworld/jw-05-step.html), mai 1998.
- [20] R. Orfali, D. Harkey, J. Edwards, *The Essential distributed Objects Survival Guide*, John Wiley, 1996.
- [21] A. Vogel, K. Duddy, *Java Programming with CORBA, advanced Techniques for Building distributed Applications*, 2nd edition, John Wiley, 1998, pp. 514.
- [22] D. Slama, J. Garbis, P. Russell, *Entreprise CORBA, Building Enterprise-Class CORBA Systems*, Prentice Hall, London, 1999, pp. 366.
- [23] R. Kimball, *The Data Warehouse Toolkit*, John Wiley, USA, 1996.
- [24] E.F. Codd, S.B. Codd, C.T. Sallley, Providing OLAP (On-Line Analytical Processing) to User Analysts: an IT mandate [on line], 1993. Disponible sur Internet : <URL:http://www.arborsoft.com/essbase/wht_ppr/coddTOC.html>.
- [25] A. Rossi Mori, et al., Exploiting the terminological approach from CEN/TC251 and GALEN to support semantic interoperability of healthcare record systems, *Int. J. Med. Inf.* 48 (1998) 111–124.
- [26] C.J. Mc Donald, et al., What is done, what is needed and what is realistic to expect from medical informatics standards, *Int. J. Med. Inf.* 48 (1998) 5–12.
- [27] B. Blobel, et al., Comparing middleware concepts for advanced healthcare system architectures, *Int. J. Med. Inf.* 46 (1997) 69–85.
- [28] B. Blobel, et al., CORBA security services for health information systems, *Int. J. Med. Inf.* 52 (1998) 29–37.
- [29] G.W. Beeler HL7 version 3: an object-oriented methodology for collaborative standards development, *Int. J. Med. Inf.* 48 (1998) 151–161.
- [30] P. Degoulet, et al., Rational and design considerations for a semantic mediator in health information systems, *Methods Inf. Med.* 37 (1998) 518–526.
- [31] W.E. Hammond, The role of standards in creating a health information infrastructure, *Int. J. Biomed. Comput.* 34 (1994) 29–44.
- [32] P. Landais, *Thésaurus de néphrologie*, Jansen Cilag Publ., Paris, 1998, pp. 99.
- [33] P. Degoulet, M. Fieschi, L'interopérabilité des systèmes d'information de santé. Aspects syntaxiques et sémantiques, CISH, CNEH édit, 1997, pp. 260.
- [34] G.P. Borand, et al., Supporting interoperability between medical knowledge-based systems: experience from pilot implementations, *Clin. Chim. Acta* 222 (1993) 23–35.
- [35] J. Dudeck, Aspects of implementing and harmonizing healthcare communication standards, *Int. J. Med. Inf.* 48 (1998) 163–171.
- [36] B. Brigl, et al., An integrated approach for a knowledge-based clinical workstation: architecture and experience, *Methods Inf. Med.* 37 (1998) 16–25.
- [37] J. Harrington, et al., The Andover Working Group: accelerating the implementation of standards, *Int. J. Med. Inf.* 48 (1998) 137–143.
- [38] J. Brender, et al., User requirements on the futur laboratory information systems, *Comput. Methods Programs Biomed.* 50 (1996) 87–93.
- [39] SIMS–REIN, Un système d'information multi-sources pour l'insuffisance rénale terminale, ministère de l'Éducation nationale, de la Recherche et de la Technologie, appel d'offres, 1999.
- [40] J.F. Sowa, *Conceptual Structures. Information Processing in Mind and Machine*, Addison-Wesley Publ., USA, 1984, pp. 481.
- [41] J.F. Sowa, Conceptual Graph Summary, in: T.E. Nagle, J.A. Nagle, L.L. Gerholz, P.W. Eklund (Eds.), *Conceptual Structures: Current Research and Practice*, Ellis Horwood Workshops, 1992.
- [42] J.F. Sowa, Conceptual analysis for knowledge-base design, *Methods Inform. Med.* 34 (1995) 165–171.
- [43] F. Volot, P. Zweigenbaum, B. Bachimont, M. Ben Saïd, J. Bouaud, M. Fieschi, J.-F. Boisvieux, Structuration on medical knowledge. Using UMLS in the conceptual graph formalism, *Proc. Annu. Symp. Comput. Med. Care* (1993) 71–113.
- [44] J.F. Smart, M. Roux, A model for medical knowledge representation application to the analysis of descriptive pathology reports, *Methods Inf. Med.* 34 (1995) 352–360.
- [45] D. Delamarre, A. Burgun, L.P. Seka, P. Le Beux, Automated coding system of patient discharge summaries using conceptual graph, *Methods Inform. Med.* 34 (1995) 345–351.
- [46] P. Zweigenbaum, B. Bachimont, J. Bouaud, J. Charlet, J.F. Boisvieux, Issues in structuring and acquisition of an ontology for medical language understanding, *Methods Inf. Med.* 34 (1995) 15–24.
- [47] P. Zweigenbaum, B. Bachimont, J. Bouaud, J. Charlet, J.F. Boisvieux, Evaluating a normalized conceptual representation produced from natural language patient discharge summaries, *Proc. AMIA Annu. Fall Symp.* (1997) 590–594.