

Etudiant : David Rapin  
Suiveur UTC : Véronique Hédou-Rouiller

UTC, GI  
Stage TN09  
Automne 2005

# Développement d'interfaces utilisateurs et de systèmes de gestion de données dans le domaine de la santé publique

Structure : Inserm  
Lieu : 27 rue Chaligny,  
75571 PARIS CEDEX 12  
Responsable : Jean-François Vibert

## Remerciements

Je remercie tout d'abord le Professeur Guy Thomas, directeur de l'unité 707, pour m'avoir accueilli au sein de son unité.

Je remercie le Professeur Antoine Flahault et le docteur Jean-François Vibert, co-responsables du réseau Sentinelles, pour m'avoir confié ce projet.

Je remercie particulièrement le docteur Thierry Blanchon, le docteur Jean-François Vibert, Clément Turbelin, Georgios Gropetis et Fabian Alvarez pour leur disponibilité, leur aide et l'intérêt qu'ils ont porté à mon stage.

J'adresse également mes remerciements au docteur Pascal Fauchart du département de santé publique de l'hôpital Tenon pour sa patience et sa collaboration.

Je souhaite signifier ma reconnaissance à tous les membres de l'unité 707 pour avoir contribué à faire de ce stage une expérience enrichissante, ainsi qu'à Valérie Lo et Alexandre Poupart, avec qui j'ai collaboré, pour leur sympathie, leur aide et leur patience.

Je remercie également Alexia Kieffer, François Fournel, Caroline Leman et Virginie Combe pour leur collaboration et leurs précieux conseils lors de la rédaction de mon rapport de stage.

Enfin, je remercie sincèrement toutes les personnes de l'Université de Technologie de Compiègne qui ont travaillé pour le bon déroulement de ce stage, et en particulier Noëlle Trigano ainsi que mon suiveur Véronique Hédou-Rouiller.

# Sommaire

|                                                                |    |
|----------------------------------------------------------------|----|
| Remerciements.....                                             | 1  |
| Sommaire .....                                                 | 2  |
| I. Introduction.....                                           | 3  |
| II. Résumé technique.....                                      | 4  |
| III. Le contexte .....                                         | 6  |
| 1. Présentation de l'Inserm.....                               | 6  |
| a. Mission.....                                                | 6  |
| b. Programmes nationaux.....                                   | 6  |
| c. Programmes européens.....                                   | 6  |
| d. Programmes internationaux.....                              | 7  |
| 2. Présentation de l'U707.....                                 | 7  |
| 3. Présentation du réseau Sentinelles et du site Sentiweb..... | 8  |
| IV. Le déroulement du stage.....                               | 9  |
| 1. Définition du sujet.....                                    | 9  |
| a. Sujets initiaux.....                                        | 9  |
| b. Sujet à l'U707 : Sentiweb.....                              | 9  |
| c. Sujet à l'hôpital Tenon : GHB++.....                        | 10 |
| 2. Le projet Sentiweb.....                                     | 10 |
| a. Identification du besoin.....                               | 10 |
| b. Répartition des tâches et travail en équipe.....            | 11 |
| c. Etude préliminaire et remise en cause du modèle de BDD..... | 11 |
| d. Planning de travail.....                                    | 13 |
| e. Détails de quelques réalisations caractéristiques.....      | 15 |
| 3. Le logiciel GHB++.....                                      | 24 |
| a. Une approche différente.....                                | 24 |
| b. Objectifs de réalisation.....                               | 24 |
| c. Organisation du travail.....                                | 25 |
| d. Détails des étapes clés de la conception.....               | 26 |
| e. Suite envisagée.....                                        | 29 |
| V. Conclusion.....                                             | 30 |
| 1. Bilan du projet Sentiweb.....                               | 30 |
| 2. Bilan du projet GHB++.....                                  | 30 |
| 3. Bilan global.....                                           | 31 |
| VI. Bibliographie.....                                         | 32 |
| VII. Annexes.....                                              | 33 |

# I. Introduction

Dans le cadre de ma formation en Génie Informatique à l'Université de Technologie de Compiègne, il m'a été demandé d'effectuer un stage en entreprise d'une durée de 24 semaines. Ce premier stage longue durée a pour but de nous confronter au monde du travail, d'en comprendre les contraintes et les enjeux, et d'assurer une première application « réelle » des enseignements que nous avons reçus jusqu'ici.

J'ai effectué ce stage du 5 septembre 2005 au 17 février 2006 à l'Institut National de la Santé Et de la Recherche Médicale, au sein de l'unité 707, dans les locaux de la faculté de médecine Saint Antoine - Pierre et Marie Curie, à Paris (XII).

Un projet parallèle m'a été confié au département de santé publique de l'hôpital Tenon, à Paris (XX), où je me suis plusieurs fois rendu pour présenter mes travaux.

A travers ce rapport, je vais rendre compte de mon travail, des difficultés rencontrées, et de la manière dont je les ai traitées. En décrivant cette démarche, je montrerai en quoi cela a sollicité mes compétences d'ingénieur.

## II. Résumé technique

D'un point de vue technique, mon stage fut particulièrement enrichissant car la réalisation de mes projets nécessita l'utilisation de nombreuses technologies.

La technologie dont ma pratique fut la plus intensive est le langage de génération de pages Web dynamiques *PHP 5*, en particulier son orientation objet qui le rend très puissant.

Le système de gestion de bases de données *MySQL 5* me fut également très utile, notamment les fonctionnalités ajoutées depuis la version 4 que je ne connaissais pas auparavant. Une importante partie de mon travail consista en l'optimisation de requêtes SQL portant sur des tables très volumineuses (plusieurs millions de lignes).

Lors de la programmation des modules *PHP* qui génèrent les cartes d'activités hebdomadaires des maladies surveillées par le réseau Sentinelles, j'ai dû interfacier les modules *PHP* avec un programme exécutable qui fournissait des résultats. J'ai été amené à recompiler ce programme en C pour le faire fonctionner sur différentes versions de Linux, ce qui a ponctuellement sollicité mes connaissances en programmation C. Ce travail m'a surtout permis d'approfondir considérablement mes connaissances des librairies graphiques de *PHP* et de leurs possibilités.

Toute une partie de mon stage a également consisté en l'installation complète d'un serveur Linux du centre de calcul de l'U707 ; serveur destiné à faire fonctionner le nouveau site Sentiweb ainsi que d'autres sites dans le futur. L'installation de ce serveur m'a donné de solides bases d'administration des systèmes d'exploitation de type Unix, en particulier des distributions de Linux *Mandrake 10* et *Red Hat Enterprise Linux 4*. De plus, j'ai acquis une certaine expérience concernant l'installation et le paramétrage sous Linux des applications *Apache*, *MySQL*, *PHP*, *vsFTPS*<sup>1</sup> et des librairies qu'elles utilisent. Aussi, la coopération de ces applications<sup>2</sup> a nécessité l'étude de la gestion des droits pour les utilisateurs et les groupes sous Linux, afin d'assurer la sécurité du serveur.

D'autres technologies ont été nécessaires au développement des applications Web que nous avons réalisées, notamment :

- *JavaScript* pour les comportements dynamiques de formulaires ;
- *CSS*<sup>3</sup> pour l'apparence et la compatibilité multi navigateurs<sup>4</sup> des pages ;
- *Smarty* comme moteur de templates<sup>5</sup> pour le site Sentiweb ;

---

<sup>1</sup> Acronyme pour Very Secure File Transfert Protocole Server. Ce logiciel est connu pour être l'application de serveur FTP la plus sécurisée sous Linux.

<sup>2</sup> La collaboration sur un serveur Linux des applications *Apache*, *MySQL* et *PHP* est connue sous le nom d'*architecture LAMP*.

<sup>3</sup> Acronyme pour Cascaded Style Sheet. Fichier de définition de la mise en forme d'un site.

<sup>4</sup> Les problèmes de compatibilité de mise en forme entre les différents navigateurs Internet sont récemment devenus préoccupants en raison du nombre croissant de personnes utilisant *Mozilla Firefox* et de son comportement très différent de celui de *Microsoft Internet Explorer*.

<sup>5</sup> Un template est un fichier définissant la structure globale d'une page - une page Web dans notre cas - et dont le contenu est ajouté dynamiquement au moment de la consultation avec un navigateur internet.

- Langage *XML*<sup>1</sup>.
- Flux *RSS*<sup>2</sup>.

Finalement, cette diversité de technologies m'a permis de cerner les enjeux du choix de ces dernières lors de la réalisation d'un projet en informatique et particulièrement lors du développement d'applications Web.

---

<sup>1</sup> Acronyme pour eXtensible Markup Language, langage extensible à balises. Le langage *XML* permet la mise en forme hiérarchique de données encapsulées dans des balises généralement de la forme « <balise>donnée</balise> ».

<sup>2</sup> Acronyme pour Really Simple Syndication. Un flux *RSS* est un fichier en langage *XML* régulièrement mis à jour et destiné à contenir un résumé formalisé du contenu récemment ajouté à un site Web. Il permet ainsi à de tierces personnes de se tenir informées des derniers changements de ce site en accédant à un seul fichier.

### **III. Le contexte**

#### **1. Présentation de l'Inserm**

##### **a. Mission**

Créé en 1964<sup>1</sup>, l'Institut National de la Santé et de la Recherche Médicale est un établissement public à caractère scientifique et technologique, placé sous la double tutelle du Ministère de la Santé et du Ministère de la Recherche.

L'Inserm est le seul organisme public de recherche français entièrement dédié à la santé humaine. Ses chercheurs ont pour vocation l'étude de toutes les maladies des plus fréquentes aux plus rares, à travers leurs travaux de recherches biologiques, médicales et en santé des populations.

La mission première de l'Inserm est de faciliter les échanges entre :

- La recherche fondamentale, définie parfois comme «guidée par la curiosité», celle dont le but est d'éclairer l'inconnu, sans but précis ;
- La recherche clinique, celle qui se fait auprès du patient ;
- La recherche thérapeutique ou diagnostique, dont le but est l'étude des maladies ;
- La recherche en santé publique, qui s'attache à mieux connaître les mécanismes de la santé par l'étude de groupes de populations notamment.

Pour remplir cette mission, l'Institut a été conçu dès l'origine dans un partenariat étroit avec les autres établissements de recherche publics ou privés, et les lieux de soins que sont les hôpitaux. Aujourd'hui, 85% des 360 laboratoires de recherche Inserm sont implantés au sein des Centres hospitalo-universitaires, ou des Centres de lutte contre le cancer, les autres étant situés sur les campus de recherche du CNRS ou encore des instituts Pasteur ou Curie.

##### **b. Programmes nationaux**

En 2004, des Programmes nationaux de recherche ou « Instituts virtuels thématiques » ont été organisés pour rassembler l'ensemble des acteurs en France dans une thématique de recherche. Ces Programmes s'inscrivent dans la durée et associent largement tous les partenaires de la recherche publique et privée, les organismes de recherche, EPST et EPIC, l'Université, les hôpitaux, les agences sanitaires, les associations caritatives et les fondations, les sociétés de spécialités médicales, les organismes de protection sociale, les associations de malades et les industriels.

##### **c. Programmes européens**

En 2004, un nouvel élan a été donné au développement d'une politique de site par delà nos frontières, soutenue à partir de projets scientifiques intégrant concrètement une dynamique d'échange de chercheurs et de définition d'objectifs conjoints dans la durée.

---

<sup>1</sup> Elle est alors le successeur de l'INH (institut national d'hygiène) fondé en 1941.

- Le programme EURYI

L'Inserm participe au Programme *EURYI* (European Young Investigators Awards), prix très élitiste porté par les responsables d'organismes de recherche européens (EUROHORCs), et coordonné par la Fondation européenne de la science (ESF). Son objectif est de soutenir les projets de jeunes chercheurs post-doctorants d'un très haut niveau scientifique, en leur offrant l'opportunité de constituer une équipe de grande visibilité internationale.

- Le 6e PCRD<sup>1</sup>

L'année 2004 a vu le développement de l'activité de la Cellule europe. Celle-ci, grâce à son étroite coordination avec la filiale privée de l'Inserm, Inserm Transfert<sup>2</sup>, a accompagné de nombreux chercheurs dans la réponse aux appels d'offres lancé par la Commission européenne dans le cadre du 6e PCRD. Ainsi, 17 projets européens sont coordonnés par des chercheurs Inserm.

Cette étroite articulation entre Inserm-Transfert et l'Inserm pour l'aide à la préparation pour le management de projets européens, permet à l'Inserm de participer à la constitution de 94 projets européens, dont il en coordonne 22 (5 réseaux d'excellence, 3 projets intégrés, 10 projets ciblés et 4 actions de soutien spécifique).

## d. Programmes internationaux

Depuis sa création, l'Inserm a mené une politique volontaire de partenariats internationaux.

Aujourd'hui, l'Inserm :

- Participe à 5412 projets de coopération avec des pays étrangers, dont 1 012 nouveaux projets mis en place en 2003 (+ 23% par rapport à 2002);
- A mis en place une cellule d'aide au montage de projets pour renforcer sa participation aux appels d'offres des NIH (Etats-unis);
- Renforce sa capacité d'accueil de longue durée de chercheurs étrangers confirmés de haut niveau par la mise en place du programme « Join Inserm ».

## 2. Présentation de l'U707

L'unité 707, intitulée « Epidémiologie, Systèmes d'Information, Modélisation », succède à l'U444<sup>3</sup> depuis le début de l'année 2005. L'unité est située dans les locaux de l'Inserm à la faculté de médecine Saint Antoine - Pierre et Marie Curie. Elle est composée de quatre équipes ayant respectivement pour thème de recherches :

<sup>1</sup> Le 6e Programme Cadre (6e PCRD) est le principal instrument communautaire de financement de la recherche en Europe pour la période 2002-2006. Il est proposé par la Commission européenne et adopté par le Conseil et le Parlement européen selon une procédure de codécision.

<sup>2</sup> Inserm-Transfert, filiale privée de l'Inserm, est une société anonyme au capital de 4 573 470 Euros. Inserm-Transfert, en collaboration étroite avec l'Inserm, contribue à la valorisation et au transfert entre recherche et industrie, notamment par la création d'entreprises. Elle facilite ainsi le développement de technologies innovantes émanant de l'Inserm.

<sup>3</sup> Les unités de recherche de l'Inserm ont une durée d'existence de 8 ans au plus. Une unité dépassant cet âge est automatiquement renommée. Ce fût récemment le cas de l'U444 qui est devenue l'U707.



- Epidémiologie des maladies infectieuses ;
- Réseau Sentinelles ;
- Les déterminants sociaux de la santé et du recours aux soins ;
- Performance hospitalière.

Ces quatre équipes se complètent et ont pour rôle commun de faire progresser les modèles mathématiques de simulation et d'étude d'épidémies, en les mettant à l'épreuve d'études sur le terrain et en milieu hospitalier.

### **3. Présentation du réseau Sentinelles et du site Sentiweb**

Le réseau Sentinelles est un système de surveillance nationale qui permet, depuis novembre 1984, le recueil, l'analyse et la redistribution en temps réel de données épidémiologiques issues de l'activité des médecins généralistes libéraux. Cette surveillance hebdomadaire concerne les pathologies transmissibles fréquentes en médecine de ville : grippe clinique, diarrhée aiguë, rougeole, oreillons, varicelle, urétrite masculine, hépatites A, B, C, asthme, tentatives de suicide et recours à l'hospitalisation.

Ce réseau fonctionne grâce à la participation bénévole de centaines de médecins généralistes répartis dans tous les départements français. Les médecins participants acceptent de passer régulièrement quelques minutes sur le site Web du réseau Sentinelles pour déclarer les cas des pathologies surveillées qu'ils ont observé récemment. La fréquence de visite demandée est de l'ordre de la semaine ou de la quinzaine.

Les déclarations de cas peuvent éventuellement être décrites sommairement à l'aide de formulaires spécifiques aux pathologies déclarées. A partir de ces déclarations de cas, des incidences<sup>1</sup> hebdomadaires estimées sont calculées pour chaque département et pour chaque région.

Les données ainsi produites permettent de surveiller l'évolution géographique et temporelle des épidémies ainsi que de prévoir l'arrivée d'épidémies à caractère saisonnier telles que la grippe ou la diarrhée aiguë. Ces données sont consultées par l'OMS, l'InVS<sup>2</sup>, ainsi que de nombreux journalistes et chercheurs en épidémiologie. Ainsi, le réseau Sentinelles est vraisemblablement le seul laboratoire de surveillance d'épidémies à rendre disponible en libre accès sur son site Web autant de données avec un tel degré de précision.

Le site Sentiweb est donc l'outil central du réseau Sentinelles, qui lui permet aussi bien de recueillir des données auprès des médecins que de traiter ces données et enfin de les présenter mises en forme.

---

<sup>1</sup> En épidémiologie, l'incidence d'une maladie dans une zone géographique sur un intervalle de temps donné est le nombre de cas de personnes infectées par la maladie. Ce chiffre se donne soit tel quel soit ramené à une population de 100 000 habitants pour le rendre plus facile à interpréter.

<sup>2</sup> L'institut de veille sanitaire (InVS), établissement public de l'Etat, placé sous la tutelle du ministère des Solidarités, de la Santé et de la Famille, a pour mission de surveiller l'état de santé de l'ensemble de la population, et d'alerter les pouvoirs publics en cas de menace pour la santé publique.

## IV. Le déroulement du stage

### 1. Définition du sujet

#### a. Sujets initiaux

Initialement, le sujet de mon stage avait été défini dans les termes suivants: « Développement du site du réseau Sentinelles corse ».

A ce sujet, dont la réalisation aurait lieu au sein de l'U707, s'ajoutait un second sujet concernant le département de santé publique de l'hôpital Tenon<sup>1</sup>. Ce second sujet était : « Conception d'un logiciel de suivi médical de patient hyperbariste<sup>2</sup> ».

#### b. Sujet à l'U707 : Sentiweb

Afin de mieux comprendre l'évolution de mon premier sujet de stage, une rapide mise en situation est nécessaire.

Durant le semestre précédant mon stage, deux stagiaires de l'université Paris VI avaient repensé et conçu une nouvelle base de données pour le réseau Sentinelles.

Les principales différences entre le nouveau schéma et l'ancien étaient une vérification de l'intégrité référentielle<sup>3</sup>, l'intégration de procédures stockées et de Triggers<sup>4</sup>, ainsi qu'une organisation des données plus dynamique.

Une grande partie de ces évolutions techniques avaient été rendues possibles par la sortie récente de la version 5 du système de gestion de bases de données *MySQL*.

Étant données les fortes dépendances structurelles entre le site Sentiweb actuel et la base de données qu'il exploite, une stagiaire de l'UTC, Valérie Lo, avait été affectée à la réécriture du site Sentiweb pour l'adapter à la nouvelle base de données.

Il est alors rapidement apparu à nos responsables et à nous-mêmes qu'il serait plus efficace en terme de travail, et plus rentable en terme de résultats de faire travailler deux

---

<sup>1</sup> Ce département est affilié à l'U707 et est sous la responsabilité d'Antoine Flahault.

<sup>2</sup> L'hyperbarie est le fait d'évoluer à des pressions supérieures à 1 bar. Cette discipline concerne principalement les plongeurs professionnels, certains ouvriers travaillant en forage et tunnels, et le personnel hospitalier travaillant en caissons de recompression.

<sup>3</sup> Lorsque des lignes d'une table d'une base de données font référence à des lignes d'une autre table, des contraintes peuvent être posées pour empêcher la suppression des lignes référencées. Bien utilisé, cet outil permet de garantir la cohérence des informations stockées.

<sup>4</sup> Les procédures stockées et les Triggers sont des fonctions dont le code source est interne au système de gestion de base de données. Pour les Triggers, l'exécution est déclenchée par un événement (par exemple la modification d'une ligne d'une table). Ces fonctions ont pour intérêt d'être transactionnelles : elles garantissent de s'exécuter de manière complète et de laisser la base de données dans un état cohérent, même en cas d'erreur.

stagiaires sur un même projet de site Sentiweb global, qui permettrait d'accéder aussi bien aux données de France métropolitaine qu'aux données corses.

## **C. Sujet à l'hôpital Tenon : GHB++<sup>1</sup>**

Mon second sujet de stage n'appelait aucune modification puisque la conception du logiciel demandé n'avait pas encore commencé et que le besoin était très précis.

Le docteur Pascal Fauchart, spécialiste du sujet et initiateur du projet d'une consultation en médecine du travail pour patients hyperbaristes à l'hôpital Tenon, m'a donné dès notre premier entretien son entière confiance pour le choix des technologies de développement et la réalisation du logiciel.

Il s'agissait donc pour moi non seulement de comprendre le besoin et de réaliser une interface répondant à la demande, mais aussi de choisir les technologies qui allaient permettre son développement.

## **2. Le projet Sentiweb**

### **a. Identification du besoin**

La première étape de mon travail en collaboration avec Valérie Lo à l'U707 fut de constituer un cahier des charges informel du site à produire en étudiant les fonctionnalités du site existant ainsi qu'en consultant les différents utilisateurs du site.

Pour un grand nombre de fonctionnalités - celles dont le fonctionnement actuel était satisfaisant - cela fut simple puisqu'il nous suffisait de reproduire l'existant en repensant le fonctionnement interne.

Cependant, la plupart des utilisateurs du produit n'étant pas informaticiens de formation, il leur était parfois difficile d'identifier les fonctionnalités qu'ils utilisaient<sup>2</sup> ou les améliorations qu'ils auraient aimées y voir apportées.

Étant donnée la situation, nous avons décidé de produire un schéma fonctionnel provisoire du site à développer (voir Annexe 1 p.34 et Annexe 2 p.35) en ayant conscience qu'un grand nombre de besoins des utilisateurs émergeraient au cours du développement et que les contenus des fonctions se préciseraient à mesure que la réalisation du site avancerait.

Le cahier des charges fut principalement constitué avec l'aide de Thierry Blanchon, médecin animateur de réseau Sentinelles, Clément Turbelin, futur médecin animateur du

---

<sup>1</sup> GHB++ est le nom donné au logiciel en projet. GHB est le sigle pour Gestion d'Hyperbaristes. Le suffixe « ++ » ou « plus plus » caractérise habituellement en informatique les logiciels non propriétaires.

<sup>2</sup> Cela est vrai en design de produit plus généralement. En effet, l'utilisateur classique n'a pas conscience du fonctionnement global du produit qu'il utilise, et n'a par conséquent pas la visibilité du champs des possibles.

réseau Sentinelles, Yves Dorléans, webmestre du site Sentiweb, ainsi que Jean-François Vibert, co-responsable de l'équipe Sentinelles.

## **b. Répartition des tâches et travail en équipe**

Une fois listées les tâches à effectuer, leur répartition se fit au fur et à mesure de la conception, de l'apparition de nouvelles fonctionnalités à ajouter, et en fonction de nos connaissances préalables ainsi que de nos intérêts personnels. La réalisation des parties suivantes me revint finalement :

- Module de cartographie spécifique à la Corse ;
- Site d'administration, gestion des déclarations ;
- Site d'administration et site Sentiweb, gestion de l'accueil dynamique et flux RSS ;
- Site d'administration, gestion des médecins ;
- Site d'administration, gestion des maladies ;
- Site Sentiweb, design du template principal (apparence et agencement général) ;
- Site Sentiweb, consultation de données (graphiques, tableaux, cartes) ;
- Installation et configuration du serveur hôte du site ;
- Manuel d'aide pour l'installation et la configuration du serveur hôte ;
- Procédure de transfert automatisé des données de l'ancienne base de données à la nouvelle ;
- Manuel d'aide à la réalisation du transfert automatisé des données ;
- Procédure de sauvegarde de la base de données ;
- Manuel d'aide à la réalisation de la sauvegarde de la base de données.

Au cours de la réalisation, Alexandre Poupart, également stagiaire de l'UTC à l'U707, aillant terminé son sujet plus tôt que prévu, se joignit à nous pour terminer les nombreuses tâches survenues au cours de la réalisation. La notion de travail en équipe fut un élément particulièrement important de notre travail sur le site Sentiweb.

En effet, travailler à deux, puis à trois a impliqué une organisation particulière. Au cours du travail, Valérie Lo, Alexandre Poupart et moi-même nous sommes régulièrement concertés pour organiser notre coopération et décider de la répartition des nouvelles tâches et corrections de bugs.

Notre manière de programmer a également été influencée par cette collaboration puisque nous nous sommes efforcés de séparer les modules du site dans des fichiers indépendants dans lesquels nous pouvions travailler simultanément sans interférer. Cette modularité a par la suite facilité la réparation de nombreux bugs en rendant trivial leur localisation dans le code.

Enfin, nous nous réunissions régulièrement avec notre responsable technique Jean-François Vibert afin de faire le point sur le travail achevé par chacun.

## **C. Etude préliminaire et remise en cause du modèle de BDD**

Une étape qui recouvrait partiellement l'étape d'identification du besoin fut la familiarisation avec la base de données conçue par les stagiaires précédents. Cette base de données était construite sous la forme de deux bases dites « transactionnelle » et « décisionnelle » (voir Annexe 3 p.36 et Annexe 4 p.37).

Dans ce type d'architecture, la première base de données sert de tampon pour l'interface de saisie et contient des données destinées à être retraitées. Son utilisation se fait principalement en écriture. La seconde est la base de lecture, dans laquelle les données validées sont écrites une fois pour toutes, puis lues pour la génération de cartes, graphiques et tableaux (voir Figure c-1).

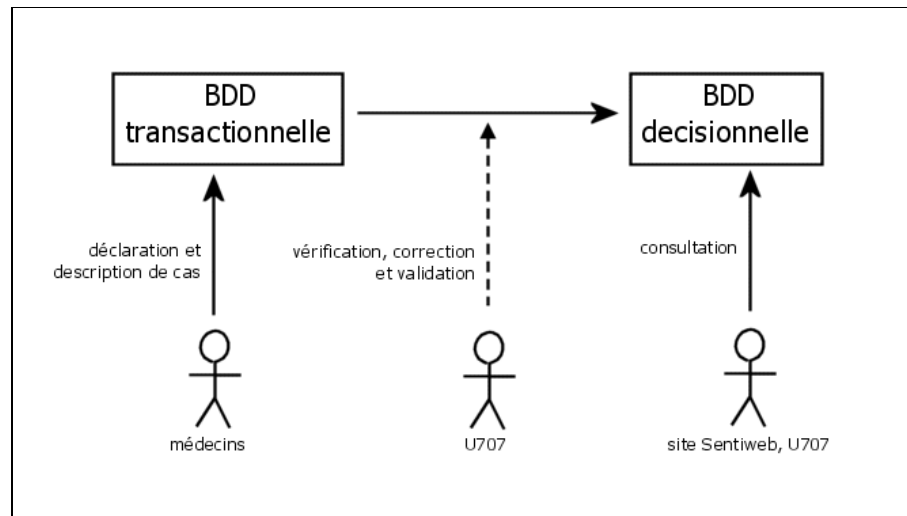


Figure c-1 : Processus de validation avec deux bases de données

Comme cela est visible sur le schéma *UML*<sup>1</sup> des deux bases de données, de nombreuses tables étaient redondantes dans les deux bases de données. Cela était lié au fait que les contraintes d'intégrité référentielle<sup>2</sup> ne pouvaient s'appliquer qu'à l'intérieur d'une même base de données.

Ainsi, par exemple, les déclarations de cas validées dans la base transactionnelle et passant dans la base décisionnelle avaient besoin d'une table de médecins et d'une table de maladies car ces deux dernières contenaient des colonnes référencées par la première. En effet, une déclaration référence le numéro du médecin qui l'a faite ainsi que le numéro de la maladie dont elle déclare le nombre de cas.

Nous avons donc constaté que, bien qu'améliorant la base de données précédente, ce modèle n'était pas satisfaisant car il contenait une redondance entraînant des risques d'incohérences d'une base de données à l'autre. De plus le double stockage entraînait des complications importantes dans la programmation de l'interface de gestion, notamment pour la gestion des maladies et des médecins.

Après avoir fait part de ces problèmes à M. Vibert, nous lui avons proposé de concevoir un nouveau schéma SQL à base de données unique se basant sur le schéma SQL

<sup>1</sup> Acronyme pour Unified Modeling Language, « langage de modélisation unifié », est un langage graphique de modélisation des données et des traitements. C'est une formalisation très aboutie et non propriétaire de la modélisation objet utilisée en génie logiciel.

<sup>2</sup> Une contrainte d'intégrité référentielle est une fonctionnalité du système de gestion de base de données. Cette fonctionnalité est conçue pour empêcher les incohérences par la méthode illustrée par l'exemple suivante. Imaginons une base de données contenant un tableau d'appartements identifiés par un numéro unique, et un tableau de locataires dont une des colonnes est le numéro de l'appartement loué. Si une contrainte d'intégrité référentielle est spécifiée pour le numéro d'appartement dans la table des locataires, un locataire ne peut être ajouté qu'avec un numéro d'appartement existant, et un appartement ne peut pas être supprimé s'il est référencé dans le tableau des locataires.

problématique. Après contrôle, ce nouveau schéma (voir Annexe 5 p.38) et son principe de fonctionnement (voir Figure e-2) furent validés.

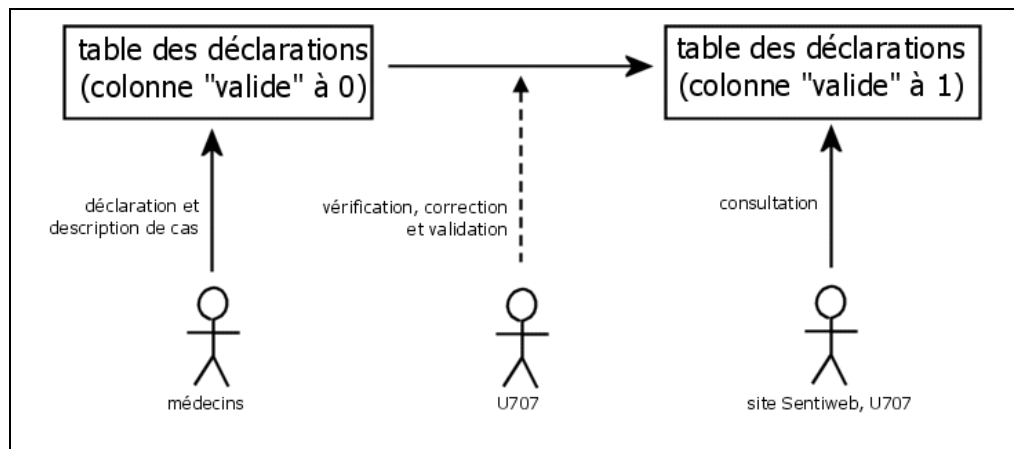


Figure c-2 : Processus de validation avec une seule base de données

Juxtaposée à cette base de données principale du réseau Sentinelles, une base de données annexe contenant les données relatives au site Web lui-même fut créée. Son contenu, non nécessaire au fonctionnement du réseau Sentinelles, fut mis à part pour améliorer la lisibilité de la première base.

La dissociation de ces deux bases était sémantiquement importante car elle évitait toute confusion entre les données nécessaires au fonctionnement du réseau et celles nécessaires au fonctionnement du site.

## d. Planning de travail

La conception du nouveau schéma de base de données nous ayant donné une visibilité suffisante pour planifier notre travail, nous pûmes proposer à nos responsables les dates butoirs suivantes :

- 20 décembre : présentation d'un site en version bêta mis en test auprès de l'équipe de l'unité.
- 15 janvier : fin de la période de test et éventuelle mise en production du site.

Ce planning fut proposé au cours de la deuxième semaine d'octobre, ce qui nous laissait plus de deux mois pour réaliser les deux sites. La fin de la période de test fut placée volontairement tôt, afin de pouvoir faire face à de très éventuels imprévus.

En pratique, la répartition des tâches dans le temps respecta l'échéancier que nous avions fixé pour le début de la période de tests, mais la période de tests fut l'occasion de la découverte de nombreuses fonctionnalités manquantes. Leur réalisation délaya sérieusement la mise en production.

La période « de sécurité » d'un mois que nous avions prévue entre la fin théorique de la période de tests et la fin de notre stage se révéla donc très utile.

La répartition effective de mes tâches dans le temps fut la suivante :

| Septembre                                                                                                                     |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
|-------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| Formalités administratives, rencontre de l'équipe                                                                             | Installation de nos postes de travail et découverte du site Sentiweb                                                                                                                    | Réécriture du module de cartographie pour la France ( <i>Mapoflash</i> <sup>1</sup> ) et écriture du module de cartographie pour la Corse                                                        | Conception et réalisation, avec Valérie Lo, du framework du site Sentiweb.                              |
| Octobre                                                                                                                       |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
| Etude de solutions et conception de la nouvelle base de données                                                               | Réalisation de la nouvelle base de données                                                                                                                                              | Réalisation du site d'administration (gestion de médecins)                                                                                                                                       | Réalisation du site d'administration (gestion déclarations)                                             |
| Novembre                                                                                                                      |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
| Design du template principal et de la feuille de styles du site Sentiweb.                                                     | Réalisation du site Sentiweb (cartes d'activité, requêtes de calcul d'incidences)                                                                                                       | Réalisation du site Sentiweb (tableaux et séries temporelles)                                                                                                                                    | Réalisation du site d'administration (gestion d'accueils dynamiques)                                    |
| Décembre                                                                                                                      |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
| Installation, configuration et sécurisation du serveur d'accueil du site.                                                     | Procédure de transfert des données de l'ancienne base à la nouvelle base.                                                                                                               | Révision du module de cartographie. Fichier de flux RSS couplé avec l'accueil dynamique. Mise en test pour l'équipe de l'U707                                                                    |                                                                                                         |
| Janvier                                                                                                                       |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
| Tests et debug. Modification de la BDD pour l'optimisation des requêtes de calcul d'incidences.                               | Tests et debug. Réalisation de parties manquantes dans l'interface d'administration (informations additionnelles pour les médecins)                                                     | Tests et debug. Procédure de sauvegarde de la base de données. Début de rédaction du rapport de stage                                                                                            | Tests et debug. Modification de la BDD pour le stockage des populations. Rédaction du rapport de stage. |
| Février                                                                                                                       |                                                                                                                                                                                         |                                                                                                                                                                                                  |                                                                                                         |
| Tests et debug. Mise en page et nettoyage du code. Début de passation aux stagiaires suivants. Rédaction du rapport de stage. | Tests et debug. Passation aux stagiaires suivants. Augmentation de la « portabilité » du code (en vue d'une installation du site sur d'autres serveurs). Rédaction du rapport de stage. | Tests et debug. Réalisation d'un outil de monitoring de l'activité des médecins sur le site. Ecriture d'une documentation de notre travail. Mise en test par des médecins du réseau Sentinelles. |                                                                                                         |

Parmi les tâches présentées ci-dessus, certaines se sont partiellement recouvertes et le planning présenté simplifie ces situations pour des raisons de lisibilité.

<sup>1</sup> De ce module, développé par un ancien stagiaire de l'UTC sous la responsabilité de Laurent Toubiana, ex membre de l'U707, n'a été reprise que la partie générant des images PNG.

## **e. Détails de quelques réalisations caractéristiques**

### **i. Réalisation de la page d'accueil de Sentiweb**

Le but de cette tâche était de réaliser un accueil modulaire, au contenu dynamique et facile à mettre à jour pour le webmestre. Le module d'accueil devait aussi être capable d'afficher l'accueil du site du réseau Sentinelles corse. Ces deux versions de l'accueil devaient être disponibles en français et en anglais.

Le modèle graphique devait rappeler l'accueil du site actuel. Une autre attente qui était une nouveauté par rapport à l'ancien site était l'archivage du contenu de la page d'accueil.

Le travail à réaliser était donc réparti en trois :

- La partie base de données et stockage des informations relatives à la page d'accueil et aux accueils archivés ;
- L'interface du site d'administration pour gérer le contenu des pages d'accueil française, corse, en français et en anglais ;
- L'affichage de l'accueil dans un template intégré au site Sentiweb.

La première étape de la réalisation de l'accueil fut de faire l'inventaire du contenu à afficher afin de mettre au clair les informations à stocker dans la base de données. Le contenu de l'accueil fut défini comme suit :

- Une date de début de validité (date de première parution de l'accueil sur le site) ;
- Un titre général pour l'accueil « France » ;
- Un titre général pour l'accueil « Corse » ;
- Une carte de France pour une maladie et une semaine donnée ;
- Une miniature de la carte de France ;
- Un titre, un sous-titre et un texte associés à la carte de France ;
- Une carte de la Corse pour une maladie et une semaine donnée ;
- Une miniature de la carte de la Corse ;
- Un titre, un sous-titre et un texte associés à la carte de la Corse ;
- Eventuellement,
  - Une image optionnelle (carte de France, carte de la Corse ou autre) ;
  - Une miniature de l'image optionnelle ;
  - Un titre, un sous-titre et un texte associés à l'image optionnelle ;
- Eventuellement,
  - Une seconde image optionnelle ;
  - Une miniature de la seconde image optionnelle ;
  - Un titre, un sous-titre et un texte associés à la seconde image optionnelle.

L'agencement graphique de ces éléments devait respecter celui de l'ancien site, dont un schéma est présenté ci-après (voir Figure e-1).



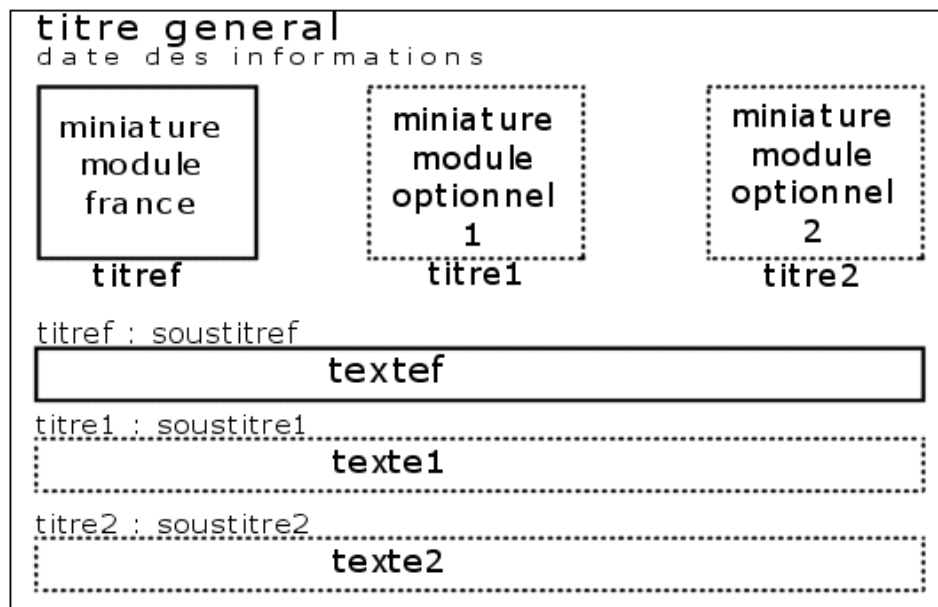


Figure e-1 : schéma du contenu de la page d'accueil

Simultanément à l'inventaire du contenu, le comportement des objets graphiques fut défini comme suit :

- Un clic sur une miniature ouvre l'image de taille normale en popup.
- En l'absence d'image pour les modules optionnels 1 ou 2, les titres, sous-titres et textes correspondants ne sont pas affichés.

L'étape suivante fut la modélisation de la structure de données qui allait stocker les informations contenues dans l'accueil dans la base de données. En l'occurrence, il fallut créer une nouvelle table dans la base de données contenant les informations structurelles du site (voir Figure e-2).

|                                        |         |
|----------------------------------------|---------|
| Identifiant                            | entier  |
| Date de début de validité              | date    |
| Titre général (France, en Français)    | texte   |
| Titre général (France, en Anglais)     | texte   |
| Titre général (Corse, en Français)     | texte   |
| Titre général (Corse, en Anglais)      | texte   |
| Affichage du module optionnel 1        | booléen |
| Affichage du module optionnel 2        | booléen |
| Titre du module France (Français)      | texte   |
| Sous-titre du module France (Français) | texte   |
| Texte du module France (Français)      | texte   |
| Titre du module France (Anglais)       | texte   |
| Sous-titre du module France (Anglais)  | texte   |
| Texte du module France (Anglais)       | texte   |
| Image miniature du module France       | image   |
| Image du module France                 | image   |
| Titre du module Corse (Français)       | texte   |
| Sous-titre du module Corse (Français)  | texte   |
| Texte du module Corse (Français)       | texte   |
| Titre du module Corse (Anglais)        | texte   |
| Sous-titre du module Corse (Anglais)   | texte   |
| Texte du module Corse (Anglais)        | texte   |
| Image miniature du module Corse        | image   |
| Image du module Corse                  | image   |
| Titre du module 1 (Français)           | texte   |
| Sous-titre du module 1 (Français)      | texte   |
| Texte du module 1 (Français)           | texte   |
| Titre du module 1 (Anglais)            | texte   |
| Sous-titre du module 1 (Anglais)       | texte   |
| Texte du module 1 (Anglais)            | texte   |
| Image miniature du module 1            | image   |
| Image du module 1                      | image   |
| Titre du module 2 (Français)           | texte   |
| Sous-titre du module 2 (Français)      | texte   |
| Texte du module 2 (Français)           | texte   |
| Titre du module 2 (Anglais)            | texte   |
| Sous-titre du module 2 (Anglais)       | texte   |
| Texte du module 2 (Anglais)            | texte   |
| Image miniature du module 2            | image   |
| Image du module 2                      | image   |

**Figure e-2 : intitulé et type des colonnes de la table "accueils"**

Une fois la table créée dans la base de données, il fallut créer la partie du site d'administration qui permet de gérer le contenu de l'accueil. Cette partie devait permettre les actions suivantes :

- Ajouter un accueil dans la base de données ;
- Modifier le contenu d'un accueil enregistré ;
- Supprimer un accueil enregistré.

Le principe de fonctionnement fut décidé comme suit :  
Le webmestre crée un accueil en donnant comme information de base sa date de début de validité. Cela signifie qu'à partir de la date de début de validité, l'accueil est affiché dans le site. Il reste affiché dans le site tant qu'aucun autre accueil avec une date de début de validité plus récente et antérieure au jour courant, n'est trouvé dans la base de données.

Figure e-3 : interface de modification du contenu d'un accueil

Le contenu textuel d'un accueil enregistré peut être modifié à l'aide d'un formulaire. En ce qui concerne les images, plusieurs cas sont possibles :

- Si le webmestre veut changer l'image relative au module France (voir Figure e-3), un bouton ouvre un popup de génération de carte de France qui lui permet de sélectionner la carte qu'il veut insérer (voir Figure e-4) ;
- Si le webmestre veut changer l'image du module corse, le principe est semblable ;
- Si le webmestre veut changer l'image d'un module optionnel, il a le choix entre y insérer une carte de France, y insérer une carte de la Corse, ou y insérer une image présente sur son ordinateur.

Figure e-4 : popup de modification de carte

Finalement, nous obtenons une page d'accueil dont le contenu est aisément modifiable, et dont le contenu complexe (cartes, images) peut être généré et enregistré en quelques clics.

Finalement, ce module est représentatif du rôle du site d'administration, qui est de permettre aux membres de l'U707 de gérer des contenus complexes à partir d'une interface simple.

## ii. Réalisation du module de gestion des déclarations

Cette partie va traiter de l'ensemble du processus de tri des déclarations, et des descriptions de déclarations. Bien que je n'aie travaillé que sur la partie relative aux déclarations, Valérie Lo ayant réalisé la partie spécifique aux descriptions, je décrirai ici l'ensemble de l'interface, ainsi que certains points de son fonctionnement interne.

La gestion des déclarations est l'étape de validation des données fournies par les médecins qui déclarent et décrivent des cas de maladies sur le site Sentiweb. Cette validation est effectuée manuellement par les membres de l'U707, et particulièrement le webmestre du site Sentiweb et le médecin animateur. Chaque personne de l'équipe qui accède à l'interface d'administration le fait avec un login d'administrateur qui permettra de l'identifier lorsqu'il triera une déclaration.

L'étape de validation manuelle permet d'assurer la pertinence des données qui sont recueillies et utilisées par la suite pour de nombreux calculs, notamment le calcul des incidences<sup>1</sup>.

Etant donné le nombre de déclarations faites par les médecins du réseau Sentinelles chaque semaine, la probabilité d'erreurs de saisie est élevée ; ainsi, il arrive que des déclarations soient faites en double ou qu'une faute de frappe s'introduise dans le nombre de cas déclarés. Il arrive également que des erreurs s'introduisent dans les réponses aux formulaires de descriptions des cas déclarés.

L'interface de gestion des déclarations, qui a été réalisée au sein du site d'administration, doit donc faciliter la détection des données aberrantes ainsi que leur correction lorsque cela est possible et jugé nécessaire. De plus, elle doit avoir une bonne ergonomie et optimiser la vitesse de travail du trieur car la tâche de tri est extrêmement répétitive.

Les données pour les déclarations et les descriptions sont principalement stockées sur quatre tables. Une description des colonnes de ces tables est disponible ci-après (voir Figure e-5 et Figure e-6).

| declarations               | declarations_valides_par_jour |
|----------------------------|-------------------------------|
| ID_DECLARATION: BIGINT(20) | ID_DECLARATION: BIGINT(20)    |
| ID_MALADIE: BIGINT(20)     | ID_MALADIE: BIGINT(20)        |
| ID_MEDECIN: BIGINT(20)     | DATE: DATE                    |
| NOMBRE_DE_CAS: SMALLINT(6) | ID_MEDECIN: BIGINT(20)        |
| DATE_DECLARATION: DATE     | NUM_INSEE: VARCHAR(6)         |
| DATE_PREMIER_CAS: DATE     | REGION: INTEGER(11)           |
| VALIDE: TINYINT(1)         | NB_CAS: DOUBLE                |
| VALIDE_PAR: VARCHAR(10)    |                               |
| EDITABLE: TINYINT(1)       |                               |

Figure e-5 : tables de stockage des déclarations

<sup>1</sup> cf. IV.2.e.iii p.22

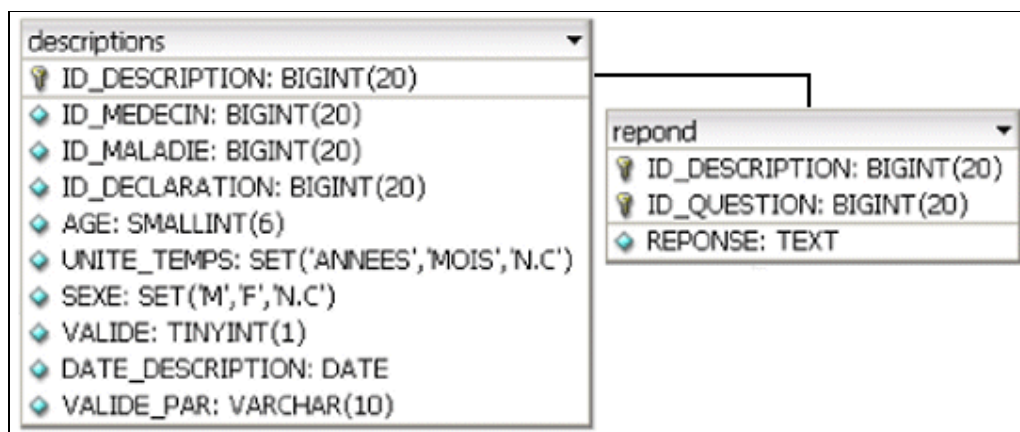


Figure e-6 : tables de stockage des descriptions

Lorsqu'un médecin fait une déclaration, par défaut, la valeur de « VALIDE » est « 0 » et celle de « VALIDE\_PAR » est « N.C ». « 0 » signifie que la déclaration n'a pas encore été triée, et « N.C » indique que l'administrateur qui a trié la déclaration est inconnu, ce qui est normal. La valeur de « EDITABLE » est initialement à « 1 » et indique que le médecin a encore la possibilité de reprendre cette déclaration pour la modifier. Pour les descriptions éventuellement liées à la déclaration, la colonne « VALIDE » est à « 0 » et « VALIDE\_PAR » est à « N.C », similairement aux déclarations.

A ce stade, la déclaration et ses descriptions associées apparaissent dans l'interface d'administration comme cela est indiqué Figure e-7.

| Declarations à contrôler. Semaine : 05 Année : 2006 . du 30/01/2006 au 05/02/2006 |            |            |                 |    |           |         |     |  |  |
|-----------------------------------------------------------------------------------|------------|------------|-----------------|----|-----------|---------|-----|--|--|
| Med.                                                                              | Decl.      | 1ier cas   | Maladies        |    |           | Valider |     |  |  |
| Lo                                                                                | 30/01/2006 | 25/01/2006 | Grippe          | 10 | (2 desc.) | 2 desc. | non |  |  |
|                                                                                   |            |            | Crises d'asthme | 11 | (1 desc.) | 1 desc. | non |  |  |

Figure e-7 : affichage de déclarations non validées, descriptions à valider

Comme cela est visible Figure e-7, la personne qui effectue le tri ne peut effectuer que deux actions sur une déclaration présentée : voir ses descriptions (bouton bleu) ou la trier comme non valide (bouton rouge). En effet le choix a été fait de ne permettre la validation d'une déclaration que si la totalité de ses descriptions associées était valide.

Si le bouton rouge est cliqué, la déclaration disparaît de l'affichage des déclarations à valider. La colonne « VALIDE » prend pour valeur « 2 » et « VALIDE\_PAR » prend pour valeur le login de l'administrateur qui a trié la déclaration.

Lorsque le bouton permettant d'afficher les descriptions est cliqué, un popup d'affichage des descriptions s'ouvre (voir Figure e-8).

| DESCRIPTIONS DE GRIPPE |      |      |                       |                 |               |                     |                         |
|------------------------|------|------|-----------------------|-----------------|---------------|---------------------|-------------------------|
|                        | Age  | Sexe | Vaccination effectuée | Hospitalisation | Si oui, motif | Si autres, précisez |                         |
| <b>Voir</b>            | 12 A | F    | O                     | O               | 3             | test lundi 30       | <b>Valider</b> <b>X</b> |
| <b>Voir</b>            | 12 A | F    | O                     | N               |               |                     |                         |

☐ Valide 
 ☐ Non valide 
 ☐ A contrôler

Figure e-8 : popup de gestion des descriptions d'une déclaration

Le popup ainsi ouvert permet de valider ou de supprimer les descriptions affichées. Le bouton « Voir » permet d'ouvrir la description dans un autre popup afin de corriger les réponses entrées. Lorsqu'une description est validée, sa colonne « VALIDE » prend la valeur « 1 ». A la fin de cette étape, chaque description de la déclaration doit avoir été éventuellement corrigée, puis validée ou supprimée. Le bouton « Terminer » (Figure e-8) permet de fermer le popup et de revenir à l'affichage des déclarations à valider (Figure e-7).

Après la validation des descriptions, la liste des déclarations apparaît comme indiqué Figure e-9. On remarquera que le bouton bleu a disparu pour laisser place à un bouton qui permet de valider la déclaration. Le trieur a alors la possibilité de modifier la valeur du nombre de cas avant de cliquer sur le bouton « oui ».

| Declarations à contrôler. Semaine : 05 Année : 2006 . du 30/01/2006 au 05/02/2006 |            |            |                              |                                        |                                    |
|-----------------------------------------------------------------------------------|------------|------------|------------------------------|----------------------------------------|------------------------------------|
| Med.                                                                              | Decl.      | 1ier cas   | Maladies                     | Valider                                |                                    |
| Lo                                                                                | 30/01/2006 | 25/01/2006 | Grippe 10 (2 desc.)          | <input type="button" value="oui"/>     | <input type="button" value="non"/> |
|                                                                                   |            |            | Crises d'asthme 11 (1 desc.) | <input type="button" value="1 desc."/> | <input type="button" value="non"/> |

Figure e-9 : affichage des déclarations non validées, descriptions validées

Lorsque la déclaration a été validée, son champ « VALIDE » prend pour valeur « 1 » et « VALIDE\_PAR » prend pour valeur le login de l'administrateur qui a trié la déclaration. Eventuellement, le champ « NOMBRE\_DE\_CAS » est modifié si la valeur a été corrigée avant validation.

Au moment où la déclaration est triée, un mécanisme interne à la base de donnée est déclenché. Il s'agit d'un Trigger, qui est une procédure stockée dans la base de donnée, écrite en script SQL, et dont l'exécution est subordonnée à un événement. Ici, c'est la modification (fonction SQL UPDATE) d'une ligne de la table « DECLARATIONS » qui déclenche l'exécution du Trigger. Le Trigger s'exécute avant l'UPDATE et compare l'ancienne et la nouvelle valeur du champ « VALIDE ». Si « VALIDE » passe de « 0 » à « 1 » (validation), la déclaration va être insérée dans la table « DECLARATIONS\_VALIDES\_PAR\_JOUR », si le champ passe de « 1 » à « 0 » (annulation du tri), la déclaration va être supprimée de la table « DECLARATIONS\_VALIDES\_PAR\_JOUR ».

La table « DECLARATIONS\_VALIDES\_PAR\_JOUR » est construite de la manière suivante : une déclaration validée contient en particulier un nombre de cas  $x$ , une date de premier cas et une date de déclaration espacées de  $y$  jours. L'hypothèse faite est que les cas déclarés sont répartis de manière équiprobable sur l'intervalle date de premier cas - date de déclaration. On fait alors  $y$  insertions dans la table « DECLARATIONS\_VALIDES\_PAR\_JOUR » pour chaque jour de l'intervalle, où chaque ligne ajoutée a un nombre de cas déclaré égal à  $x/y$ .

L'interface de gestion permet aussi d'afficher les déclarations déjà triées, valides et invalides, et de consulter leurs descriptions associées. Ces déclarations peuvent repasser dans l'état non trié ; seul le super administrateur du site, dont le login est « superadmin », détient ce pouvoir.

Finalement, cette partie particulière de notre travail, faisant le lien entre l'interface des médecins et la gestion des données recueillies, nous a demandé un effort de travail tout particulier. En effet elle demandait l'assurance d'une fiabilité sans faille et d'une ergonomie avancée, le tout reposant sur un fonctionnement technique complexe. Par les différents problèmes techniques qu'elle a soulevés, et les choix que nous avons dû faire pour accorder les contraintes techniques aux nécessités ergonomiques, cette réalisation fut extrêmement intéressante.

### **iii. Optimisation du calcul des incidences**

L'incidence d'une maladie dans une zone géographique et sur un intervalle de temps défini est le nombre de personnes infectées par cette maladie. Les incidences sont les principales données produites par le réseau Sentinelles ; ce sont les incidences hebdomadaires départementales, régionales et nationales des maladies surveillées par le réseau. Estimées à partir des données recueillies auprès des médecins, elles sont traduites en cartes d'activité, séries temporelles et tableaux de données disponibles sur le site Sentiweb.

La méthode pour estimer les incidences d'une zone géographique sur un intervalle de temps nécessite la connaissance de trois paramètres (voir Figure e-11) :

- Le cumul de la participation journalière des médecins du réseau Sentinelles habitant cette zone géographique (dans l'intervalle de temps donné);
- Le nombre de médecins en activité dans la zone géographique à ce moment;
- Le cumul du nombre des cas déclarés pour cette maladie par les médecins du réseau Sentinelles habitant cette zone géographique (dans l'intervalle de temps donné).

La participation journalière est le nombre de médecins différents qui ont fourni des données, mêmes nulles, pour une journée. Un exemple de requête SQL de calcul de la participation pour la semaine 2005-40 dans la région 21 (Île-de-France) est donné ci-dessous (voir Figure e-10).

```
SELECT
    count( distinct ID_MEDECIN, DATE ) as participation
FROM
    DECLARATIONS_VALIDES_PAR_JOUR
WHERE
    ( yearweek( DATE , 3) = 200540 )
AND
    ( REGION = 21 )
;
```

**Figure e-10 : exemple de calcul de participation en langage SQL**

Le cumul de participation ainsi que le nombre de cas déclarés sont calculés à partir de la table « DECLARATIONS\_VALIDES\_PAR\_JOUR » (voir Figure e-5 p.19). Le nombre de médecins en activité dans une zone géographique est stocké dans la base de données et est mis à jour annuellement grâce à des données fournies par le Conseil National de l'Ordre des Médecins.

Soient :

- M une maladie
- Z une zone géographique
- T un intervalle de temps

$$\text{Incidence (M, Z, T)} = \frac{\text{NombreDeCas (M, Z, T)} * \text{NombreDeMedecin (Z, T)}}{\text{Participation (Z, T)}}$$

Figure e-11 : Le calcul d'une incidence

Une forte contrainte technique, qui est apparue durant les périodes de tests de l'application, est la lenteur effective du calcul des cumuls de participation et du nombre de cas. Cette lenteur étant exponentiellement proportionnelle à la longueur de l'intervalle de temps pour lequel le cumul est calculé. Le délai de calcul, pouvant dépasser 60 secondes, était dû au très grand volume de la table « DECLARATIONS\_VALIDES\_PAR\_JOUR ». Cela était particulièrement gênant dans la mesure où il était prévu que ce calcul soit fait au moment de la génération de cartes d'activité, séries temporelles et tableaux de données.

Après avoir tenté de nombreuses méthodes d'optimisation de ces calculs, il est apparu que nous ne pouvions pas raisonnablement garder ce principe de fonctionnement, qui mettait en danger les performances du serveur Web, dont la disponibilité ne devait en aucun cas être affectée.

Finalement, la méthode retenue fut de pré calculer hebdomadairement les incidences à un moment de faible charge du serveur, et de les stocker dans la base de données. Un problème lié à cette méthode est qu'un médecin est autorisé à faire des déclarations dont la date de premier cas remonte à quinze jours (au plus). Cela implique que le nombre de cas déclarés dans les quinze derniers jours peut changer, et donc que les incidences qui avaient été calculées pour les semaines contenant ces jours peuvent changer (voir Figure e-12).

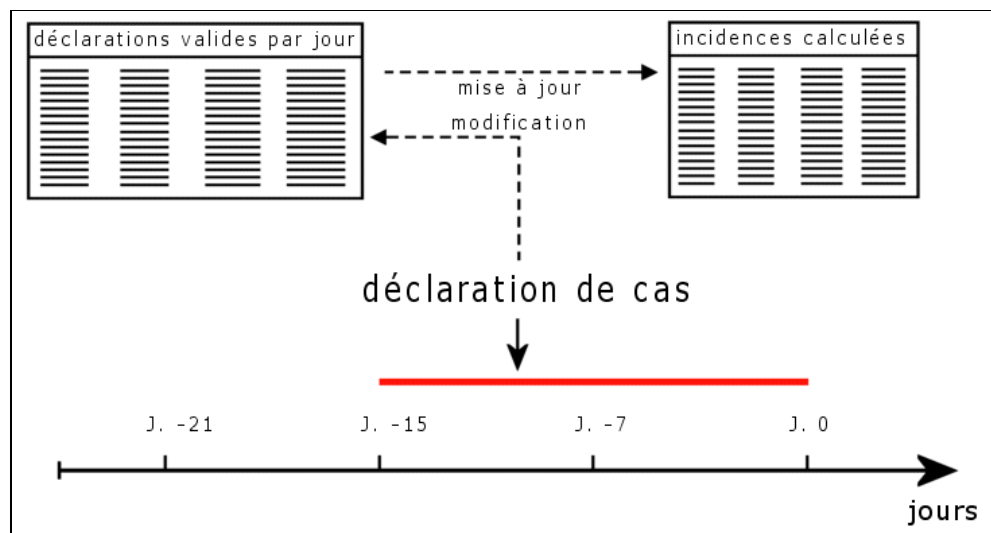


Figure e-12 : conséquences d'une déclaration avec date de premier cas éloignée

Nous constatons qu'une déclaration peut donc modifier les incidences jusqu'à trois semaines en arrière. En effet, si J. 0 est un dimanche, J. -15 est le dimanche de la semaine S. -3. Une procédure a donc été conçue pour recalculer chaque lundi matin à 3h les incidences hebdomadaires des trois dernières semaines ; et ce pour chaque maladie, région et département.



Grâce à cette solution, le serveur ne fait plus aucun calcul d'incidence au moment de l'affichage de cartes, séries temporelles et tableaux de données. Il gagne ainsi en temps de réponse et évite de recalculer inutilement les incidences éloignées de plus de trois semaines, qui ne changeront plus.

Cette partie de mon travail illustre la démarche d'optimisation du temps d'exécution de requêtes à la base de données. En effet, la vitesse d'exécution de la majorité des requêtes est liée au nombre d'enregistrements dans la table sur laquelle porte la requête. Cela rend la détection des requêtes lentes difficile et nécessite leur test dans des conditions semblables aux conditions de production. Les solutions qui sont trouvées pour améliorer les temps de calcul doivent elles-mêmes être testées afin d'assurer que leur fonctionnement ne sera pas ralenti par d'autres facteurs.

Finalement, les solutions techniques apportées doivent prendre en compte l'évolution future de la base de données et nécessitent une bonne compréhension du fonctionnement interne du SGBD<sup>1</sup> pour comprendre et prévoir les éventuels problèmes.

### **3. Le logiciel GHB++**

#### **a. Une approche différente**

Mon travail lors de la réalisation du logiciel GHB++ fut influencé par le projet Sentiweb. En effet, les premières phases de sa réalisation se firent plus tard que pour mon projet à l'U707. Ainsi, mon travail et ma réflexion sur l'organisation de ce dernier se firent avec l'expérience acquise au cours du premier projet.

J'avais donc très présente à l'esprit l'importance de la constitution d'un cahier des charges précis au moment de mes rencontres avec Pascal Fauchart ; responsable du projet intégrant le logiciel GHB++. La contrainte de ne pouvoir rencontrer mon responsable de projet que ponctuellement m'obligea aussi à formaliser le besoin afin d'être plus autonome dans mon travail.

#### **b. Objectifs de réalisation**

L'application à réaliser devait permettre de gérer des dossiers de patients hyperbaristes dans un contexte de médecine du travail. Cela impliquait un suivi très complet des patients en raison des risques de leur métier. Une juridiction très précise existe à ce sujet, définissant les points à contrôler par les médecins du travail.

Le but était donc de permettre aux médecins de stocker pour les patients suivis une fiche d'informations personnelles, l'état des vaccinations, ainsi que le contenu des visites médicales effectuées.

Chaque visite médicale est constituée de modules regroupant un certain nombre de points de contrôle. A une visite enregistrée, il devait être possible d'effectuer des ajouts, contenant des mises à jour concernant un nombre défini de modules.

---

<sup>1</sup> Acronyme pour Système de Gestion de Base de Données. Dans notre cas, le SGBD est MySQL 5.

Les éléments principaux de l'interface étaient donc la saisie des informations personnelles, et la saisie des visites médicales. Un accent particulier devait être mis sur la sécurité, en raison de la sensibilité des données médicales enregistrées. Enfin, l'interface devait par sa simplicité et son intuitivité, rendre facilement accessible et compréhensible le grand nombre de données stocké.

## C. Organisation du travail

La réalisation du logiciel, qui n'est à ce jour pas encore terminée, peut être découpée en six étapes principales :

- Formulation du besoin ;
- Choix des technologies ;
- Réalisation du cahier des charges ;
- Réalisation du framework<sup>1</sup> et du schéma de base de données sous-jacent ;
- Réalisation du corps de l'application (formulaires, interface) ;
- Mise en test, correction et finition.

En pratique, les deux premières étapes eurent lieu lors de la première réunion avec Pascal Fauchart, Antoine Flahault et Jean-François Vibert, le premier jour de mon stage. D'après le besoin formulé par Pascal Fauchart, je proposai l'utilisation de *EasyPHP*<sup>2</sup> et la réalisation du logiciel comme site Web local et privé<sup>3</sup>. Cette proposition fut validée par Jean-François Vibert et Antoine Flahault.

La troisième étape se déroula sous la forme de trois réunions avec Pascal Fauchart où je pus l'interroger très précisément sur les fonctionnalités désirées et ses attentes diverses. Suite à ces réunions, un schéma de principe de l'application fut proposé, et validé (voir Annexe 7 p.40). Ce schéma, en plus de récapituler les besoins, fixe aussi un principe de fonctionnement, dont le choix détermine entièrement la manière dont l'application est conçue en interne.

La quatrième étape fut la plus intéressante, car énormément de solutions s'offraient à moi. Ainsi, les décisions prises allaient déterminer toute la suite de mon travail, et la robustesse de l'application. C'est aussi l'intérêt personnel très fort que je porte à cette phase de la conception qui m'a poussé à approfondir cette étape.

La cinquième étape fut la plus longue car la réalisation des formulaires de saisie, qui représentent la majeure partie de l'interface, fut particulièrement longue et souleva des problèmes non entrevus lors de la conception de la logique de fonctionnement et du framework.

---

<sup>1</sup> Le framework d'une application est son squelette logique, il définit son fonctionnement global et l'organisation de ses parties les unes par rapport aux autres.

<sup>2</sup> *EasyPHP* est un logiciel pour *Microsoft Windows* qui fait fonctionner un serveur Web *Apache* avec un module *PHP*, ainsi qu'un serveur de base de données *MySQL*. Ce logiciel est très pratique dans la mesure où son installation est extrêmement simple, mettant ces services à la portée d'utilisateurs non expérimentés.

<sup>3</sup> Un site Web local est un site Web stocké sur la machine de l'utilisateur unique du site. L'utilisateur accède au site avec un navigateur Web classique via l'adresse `http://127.0.0.1` qui représente son propre ordinateur. Ce site Web est privé si seul l'ordinateur qui l'héberge peut y accéder.

Enfin, la dernière étape fut de soumettre le travail réalisé à Pascal Fauchart afin qu'il détecte les anomalies et les éventuelles fonctionnalités manquantes. Cette phase est encore en cours et il est question que je sois chargé de la terminer en dehors du contexte de ce stage.

## d. Détails des étapes clés de la conception

### i. Conception du framework

La réalisation du framework fut principalement une phase de réflexion sur la manière dont les différentes fonctions du logiciel allaient être agencées. Je choisis finalement une architecture reposant sur des contextes. Ce principe permet de ne mettre à la portée de l'utilisateur que les fonctionnalités qui lui sont nécessaires dans le contexte courant, et ainsi de simplifier l'interface (voir Annexe 8 p.41).

Avant de commencer la programmation du corps du logiciel, il me fallait concevoir un modèle logique comme base. La version de *PHP* utilisée dans *EasyPHP* étant orientée objet, je pus réaliser un modèle logique sous forme de classes (voir Annexe 9 p.42). Un objectif de ce modèle logique était de construire une couche d'abstraction rendant transparente les interactions avec la base de données.

L'utilisation de tout le potentiel de la programmation Web orientée objet m'a permis de réduire considérablement le volume du code source dont le nombre de lignes ne dépasse pas 4500, avec moins de 1400 pour le cœur du programme : les définitions de classes (voir Figure d-1).

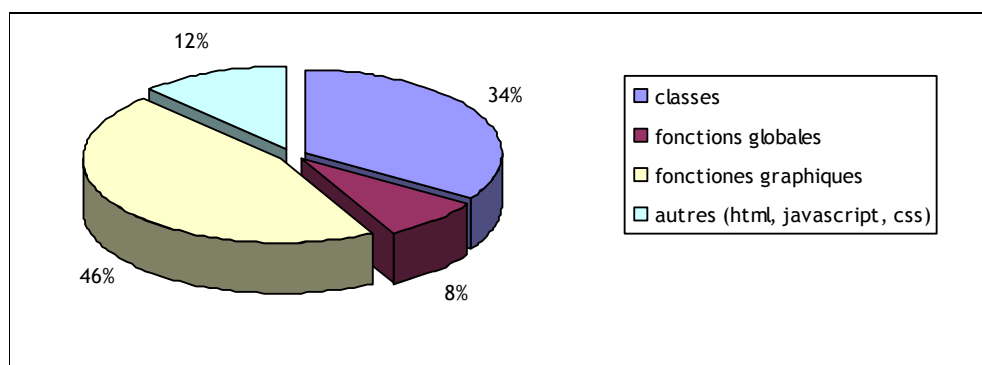


Figure d-1 : Répartition des lignes du code source de GHB++

Le framework a été axé autour de l'importance des données entrées et de leur protection. En effet lors de la navigation entre les modules d'une visite, chaque module quitté est instantanément sauvegardé dans la base de données. De cette manière, un arrêt violent de l'application (volontaire ou non) n'entraîne pas la perte des données. Lors du démarrage de l'application, le système recherche des visites médicales non clôturées et propose de reprendre la saisie à l'endroit où elle s'était arrêtée. Cette fonctionnalité est très importante dans la mesure où une visite médicale contient un très grand nombre de données dont la saisie peut représenter plusieurs heures de travail.

## ii. Conception et réalisation de base de données sous-jacente

La base de données sous-jacente à l'application fut réalisée en étroite collaboration avec Pascal Fauchart, car toutes les questions d'une visite médicale de plusieurs heures ont dû être formalisées, et leurs réponses possibles énumérées.

Une fois ce travail de formalisation des données effectué, il fallut construire une base de données avec les contraintes suivantes :

- Créer pour chaque patient un identifiant unique et persistant, permettant d'éviter les collisions, même au niveau national, dans l'hypothèse de plusieurs médecins échangeant des dossiers patient ;
- Créer pour chaque visite médicale un identifiant unique et persistant pour les mêmes raisons ;
- Structurer les données en vue d'une exportation facile vers un fichier, pour l'échange de dossiers patient entre médecins ;
- Sécuriser l'accès aux données dans *MySQL* pour empêcher leur accès sans mot de passe, même par d'autres logiciels.

Le schéma final (voir Annexe 10 p.43) répond ainsi contraintes imposées :

- Chaque patient a un identifiant généré à partir de ses initiales, de sa date de naissance, plus quatre chiffres tirés au hasard. Cela rend pratiquement impossible la génération de plusieurs identifiants identiques.
- L'identifiant d'une visite médicale est généré à partir de la date à la seconde près de sa création, plus deux chiffres tirés au hasard. Cela rend également très improbable l'existence d'identifiants identiques.
- Les données sont structurées avec un référencement de profondeur 3 au maximum, on peut ainsi exporter toutes les données d'un patient *x* avec une double boucle imbriquée selon l'algorithme suivant : exporter toutes les lignes de tables qui référencent le patient *x*, pour chaque visite *y* référençant *x*, exporter toutes les lignes des tables de modules référençant la visite *y*.
- Les login et mot de passe d'accès au logiciel sont simultanément ceux d'accès à la base de données. Après l'installation, l'utilisateur de GHB++ est aussi le seul utilisateur *MySQL*. Ainsi les données ne sont jamais accessibles sans mot de passe, même lors d'un accès manuel à la base de données.

Un enregistrement de la CIM10<sup>1</sup> était également nécessaire dans la base de données, afin de codifier les antécédents médicaux personnels et familiaux des patients. Cela impliqua un travail de recherche et de compréhension de la classification, et le développement d'un petit outil permettant sa navigation (voir Figure d-2).

---

<sup>1</sup> Classification Internationale des Maladies, version 10. Cette classification sous forme d'arborescence à trois étages (chapitres, blocs et pathologies) regroupe plus de 30 000 pathologies et les représente par un code de 3 à 6 lettres.

http://192.168.0.100 - GHB++ - Recherche CIM10 - Mozilla Firefox

Selection du code CIM10 pour le diagnostic :

(F00-F99) Chapitre V Troubles mentaux et du comportement

(F90-F98) Troubles du comportement et troubles émotionnels apparaissant habituellement durant l'enfance et l'adolescence

(F91) Troubles (de) (des) envahissants du développement

(F91.2) Ecole buissonnière

selectionner

selectionner

Selection :

|             |                    |
|-------------|--------------------|
| Code CIM10: | F91.2              |
| Intitulé:   | Ecole buissonnière |

annuler

confirmer la selection

Figure d-2 : Outil de navigation de la CIM10 intégré à l'application

### iii. Réalisation du corps de l'application

Cette étape, la plus importante en terme de temps, fut quasi-totalement consacrée à la création des fonctions graphiques. Ces fonctions sont destinées à afficher l'interface de l'application et principalement les formulaires.

Pour les formulaires des modules de visite, la lecture et la sauvegarde du contenu dans la base de données est totalement gérée par la couche objet créée lors de la réalisation du framework. Les fonctions d'affichage ont été conçues suffisamment génériques pour pouvoir afficher les données en lecture seule (consultation d'une visite) ou en édition (saisie d'une visite), afin d'en limiter le nombre.

Un moment problématique de la réalisation du corps de l'application fut la découverte d'un cas non traité par la couche objet. Il s'agit du cas de modules pour lesquels il existe plusieurs entrées dans la base de données pour une même visite. Comme cela est illustré Figure d-3, une visite peut demander l'enregistrement des trois accidents en hyperbarie ayant eu lieu depuis la visite précédente. Cela diffère des autres modules comme par exemple le module cardiologique, qui ne contient qu'un seul diagnostic pour une visite donnée.

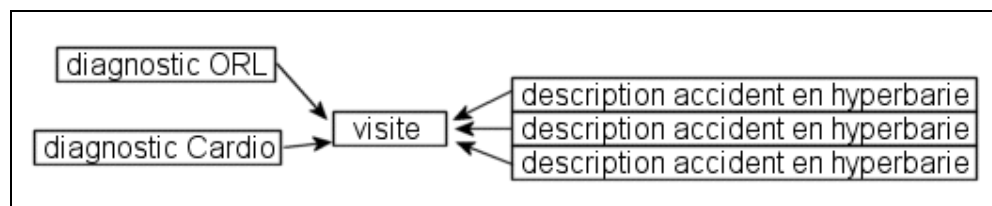


Figure d-3 : exemple d'un module ayant plusieurs entrée pour la même visite

Ce cas de figure n'avait pas été envisagé lors de la création de la classe « Module » (voir Annexe 9 p.42) qui ne permettait la modification et l'enregistrement que d'une seule ligne de la base de données par visite et par module. Des fonctions supplémentaires ont donc dû être créées dans cette classe, afin de gérer le cas particulier où un module contenait une liste d'éléments pour une visite donnée.

Finalement, les problèmes survenus lors de la réalisation du corps de l'application furent rares et assez facilement résolus, grâce au soin que j'avais apporté à la conception du framework.

## **e. Suite envisagée**

A ce jour, l'application est pleinement opérationnelle et stable, mais l'intégration de certaines fonctionnalités manque encore. En l'occurrence, la possibilité d'imprimer en une seule fois toute une visite médicale en format papier doit être ajoutée.

De plus, un problème de confidentialité lors de l'exportation du dossier patient doit être réglé. En effet, la CNIL<sup>1</sup> impose que, lors du transfert informatique d'un dossier médical, les données médicales ne doivent pas être jointes aux données permettant l'identification personnelle du patient<sup>2</sup>. Ainsi l'exportation du dossier patient sous forme de fichier doit être modifiée pour se soumettre à cette contrainte.

Enfin, une interface d'installation guidée doit être développée afin de permettre à des personnes sans formation informatique d'installer et d'utiliser l'application sur leur ordinateur personnel. Cette interface sera vraisemblablement développée avec l'outil gratuit de création de paquets d'installation *NSIS*<sup>3</sup>.

Il est question que je termine l'application GHB++ en réalisant ces modifications finales en dehors du contexte de ce stage. Si cela ne devait pas être le cas, le programme et son code source ont été clairement documentés pour permettre à une tierce personne de faire ce travail aisément.

---

<sup>1</sup> Commission Nationale Informatique et Liberté. Commission française chargée de la protection des données personnelles traitées et stockées de manière informatique.

<sup>2</sup> Source : « Informatisation du dossier médical de médecine du travail et épidémiologie : pourquoi et comment », M. Goldberg, E. Imbernon.

<sup>3</sup> Nullsoft Scriptable Install System, <http://nsis.sourceforge.net> .

## **V. Conclusion**

### **1. Bilan du projet Sentiweb**

Le projet de refonte du site Sentiweb m'a énormément apporté dans ma conception du travail en équipe et de son organisation. Grâce à la confiance qui nous a été faite, Valérie Lo, Alexandre Poupart et moi-même avons appris à organiser notre coopération dans la réalisation d'un projet qui dépassait vraisemblablement les compétences de chacun d'entre nous.

Les multiples aspects de ce projet en ont aussi fait pour moi le moyen d'élargir et d'approfondir ma culture technique. La forte pression résultant du fait que nous travaillions sur un site qui allait effectivement être mis en production et recueillir des données nationalement utilisées a été très motivante. L'utilisation des interfaces produites par des non informaticiens nous a également imposé un niveau de finition que nous n'avons jamais l'occasion d'atteindre lors de projets d'exercice dans le cadre de nos cours.

D'autre part la reprise de notre travail par d'autres stagiaires nous a imposé de documenter notre travail et d'effectuer une passation. Cette démarche de synthèse de notre travail et de transfert de connaissances fut très enrichissante.

Enfin, ce travail au sein d'une unité de recherche était pour moi une démarche visant à connaître mes affinités avec le milieu de la recherche. Je considère que cette expérience a été concluante, et envisage sérieusement de m'orienter professionnellement dans cette direction.

### **2. Bilan du projet GHB++**

La réalisation complète de cette application, en particulier avec l'expérience récemment acquise dans le cadre du projet Sentiweb, m'a permis de comprendre l'importance des différentes étapes de la chaîne de production d'une application. J'ai également pu mettre en application des connaissances en design de produit acquises au cours de ma formation.

Aussi, la nécessité de comprendre un besoin exprimé par une personne non informaticienne et la rédaction d'un cahier des charges à partir de cette collaboration ont constitué des étapes jamais abordées dans le cadre scolaire. Cela a donné pour moi un éclairage nouveau au métier d'ingénieur en informatique, qui est de comprendre et de formaliser le besoin.

Enfin, les problématiques abordées au cours de la réalisation de ce projet ont confirmé mon intérêt pour l'étude de l'ergonomie des logiciels, sujet que je mettrais au centre de ma formation dans la filière ICSI à l'UTC.

### 3. Bilan global

La réalisation de ces deux projets a finalement été pour moi l'occasion de découvrir d'autres aspects de ce que nous apprenons dans le cadre scolaire, plus qu'une simple application de mes connaissances.

Le cadre de mon stage n'ayant pas été un milieu majoritairement informaticien, j'ai du apprendre à écouter et à interpréter les requêtes des personnes, et à les traduire en produits. Aussi, certaines situations m'ont amené à justifier des choix techniques complexes sans utiliser de termes techniques. Cela m'a poussé à raisonner de manière plus abstraite, et finalement, m'a moi-même permis de mieux comprendre les enjeux de mon travail.

Aussi, la grande autonomie que j'ai eue dans mes deux projets m'a fortement responsabilisé et m'a appris à prendre le temps de mesurer les conséquences à long terme de mes choix techniques. Cela m'a révélé un aspect de la démarche de conception que je n'avais jamais soulevé dans le cadre scolaire.

Finalement, les aspects techniques et humains de mon stage m'ont aidé à comprendre le rôle d'un l'ingénieur en informatique dans une entreprise. Ce rôle est de faire le lien entre les personnes qui expriment un besoin et la solution technique adaptée à ce besoin. Faire ce lien nécessite une forte aptitude à communiquer, une bonne connaissance des différentes solutions possibles et la capacité à faire des compromis pour accorder la solution au besoin. Ce stage, en me mettant dans ce rôle particulier, m'a appris à raisonner comme un ingénieur.



## VI. Bibliographie

Documentation *PHP* :

<http://fr2.php.net/manual/fr/>

Documentation *MySQL* :

<http://dev.mysql.com/doc/refman/5.0/en/>

Documentation *Smarty* :

<http://smarty.php.net/manual/en/>

Documentation *JavaScript* et *CSS* :

<http://www.quirksmode.org/>

Documentation flux *RSS* :

<http://blogs.law.harvard.edu/tech/rss>

Informatisation du dossier médical :

« Informatisation du dossier médical en médecine du travail et épidémiologie : pourquoi et comment ? », M. Goldberg et E. Imbernon, Masson, Paris, 1994.

Sources *CIM10* :

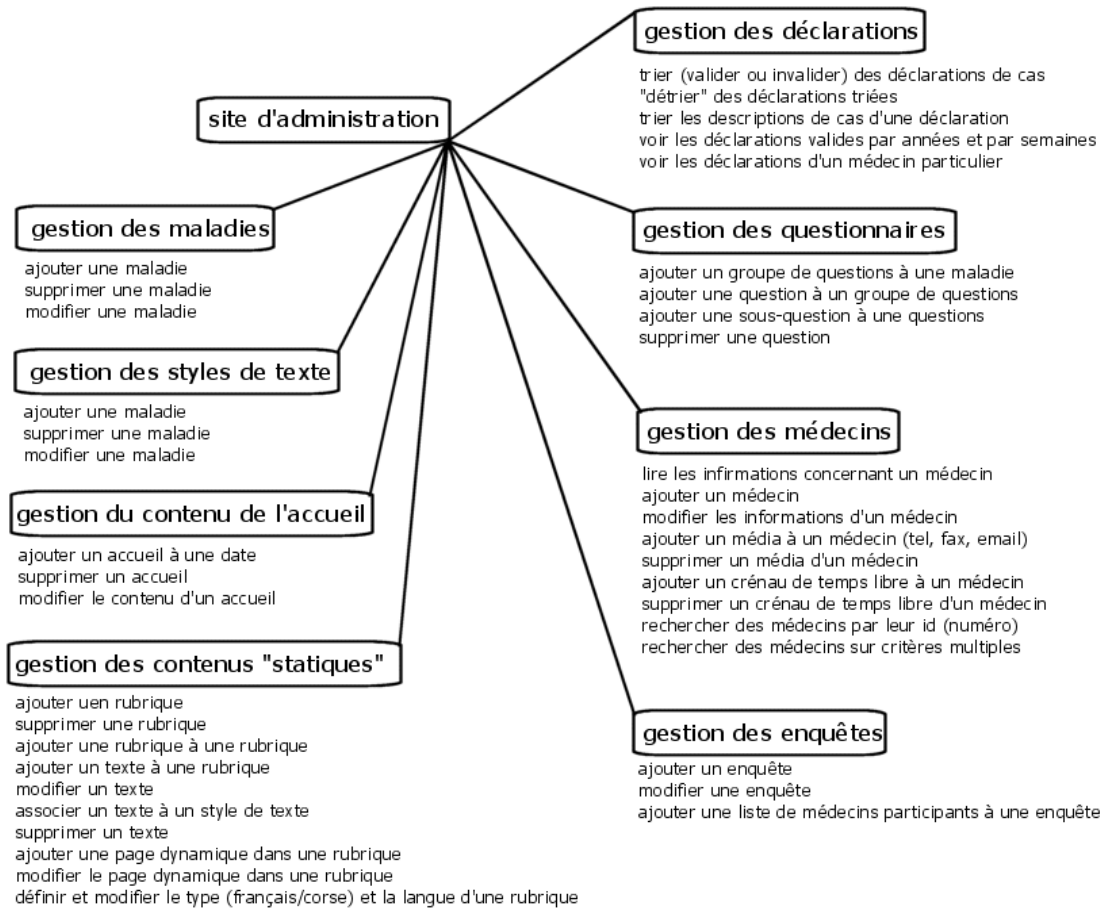
<http://taurus.unine.ch:8080/icd10>

« Définition en entités et relations de la CIM10, spécifications d'implémentation », Office Fédéral de la Statistique, Lausanne, 2004.

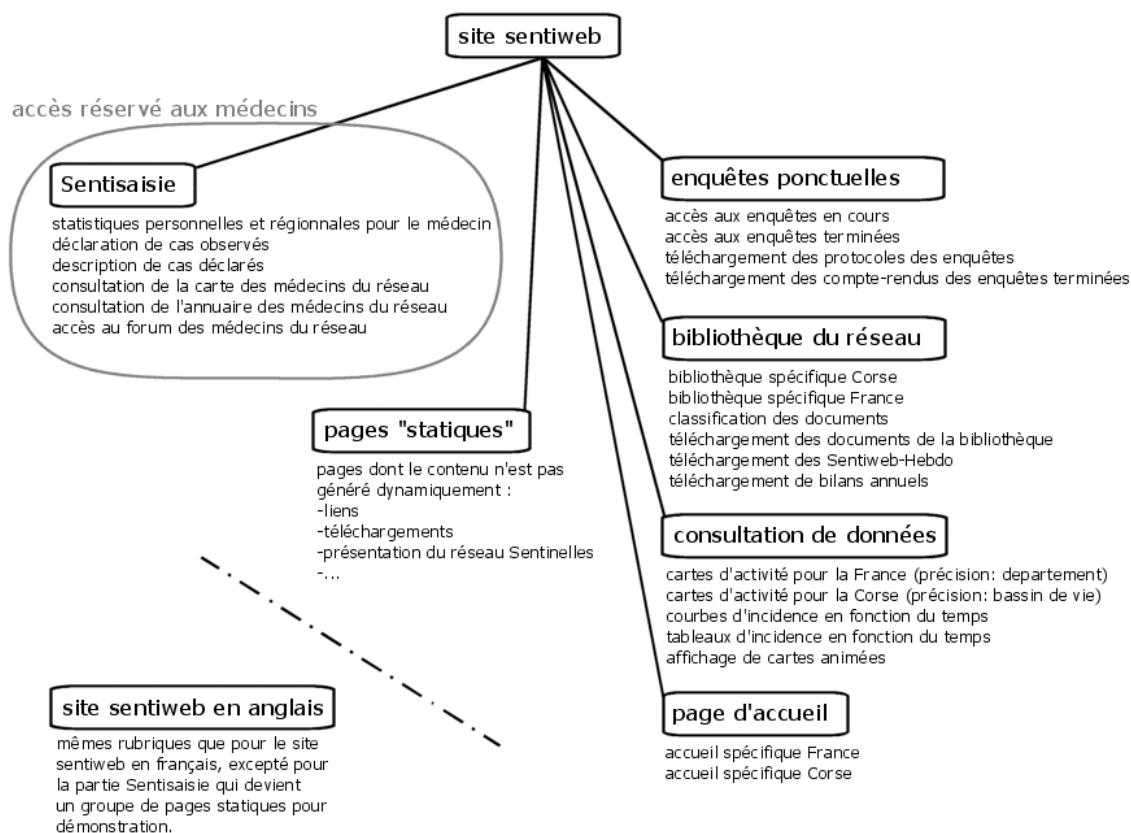
## VII. Annexes

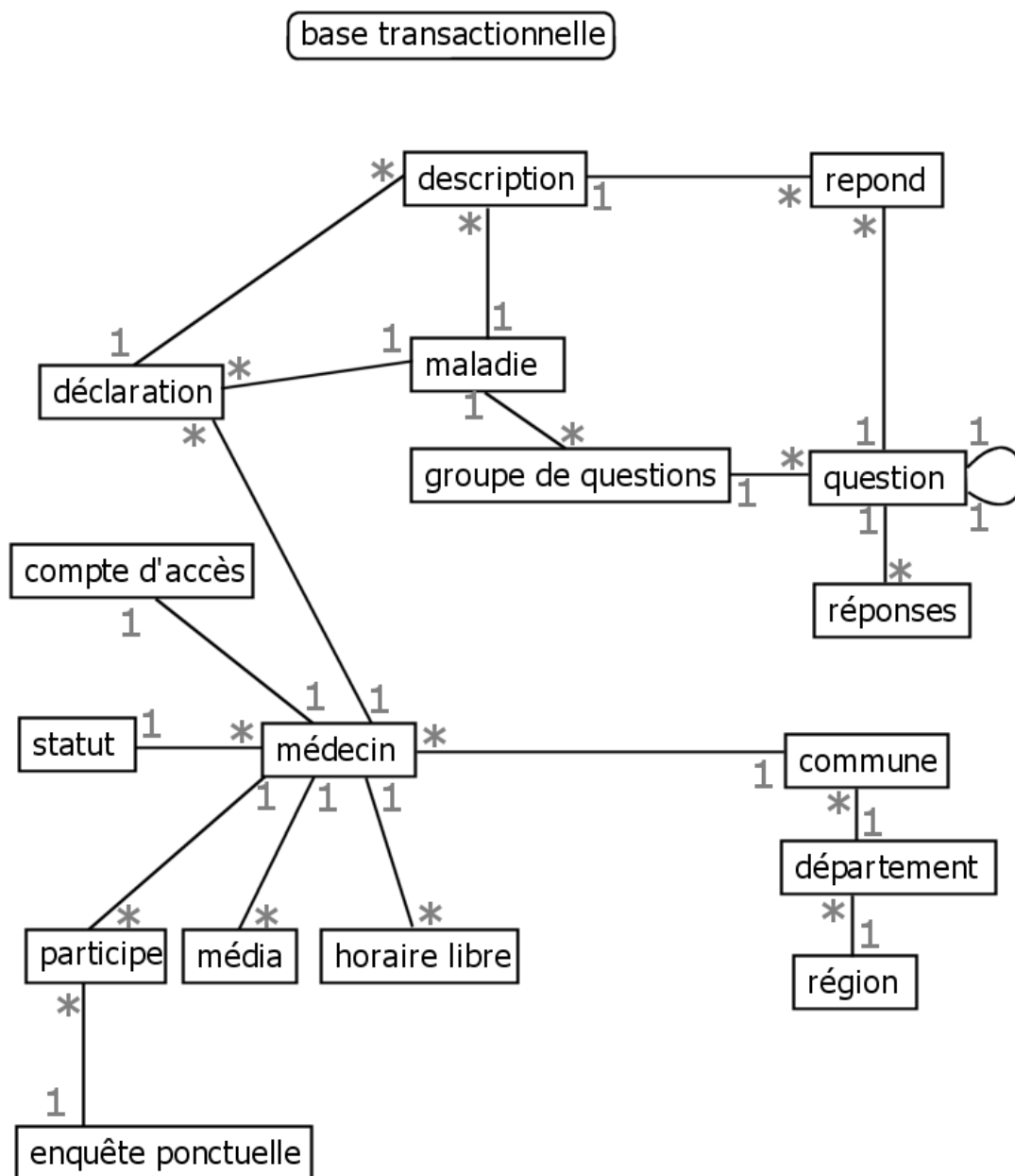
|                                                                            |    |
|----------------------------------------------------------------------------|----|
| Annexe 1 : Diagramme des fonctions du site d'administration .....          | 34 |
| Annexe 2 : Diagramme des fonctions du site Sentiweb .....                  | 35 |
| Annexe 3 : Diagramme UML simplifié de la BDD transactionnelle .....        | 36 |
| Annexe 4 : Diagramme UML simplifié de la BDD décisionnelle.....            | 37 |
| Annexe 5 : Diagramme UML simplifié de la BDD Sentinelles unique .....      | 38 |
| Annexe 6 : Schéma complet de la BDD Sentinelles .....                      | 39 |
| Annexe 7 : Schéma fonctionnel de l'application GHB++ .....                 | 40 |
| Annexe 8 : Contextes et fonctions relatives dans l'application GHB++ ..... | 41 |
| Annexe 9 : Schéma UML simplifié des classes de l'application GHB++.....    | 42 |
| Annexe 10 : Schéma UML de la BDD sous-jacente à l'application GHB++ .....  | 43 |

## Annexe 1 : Diagramme des fonctions du site d'administration

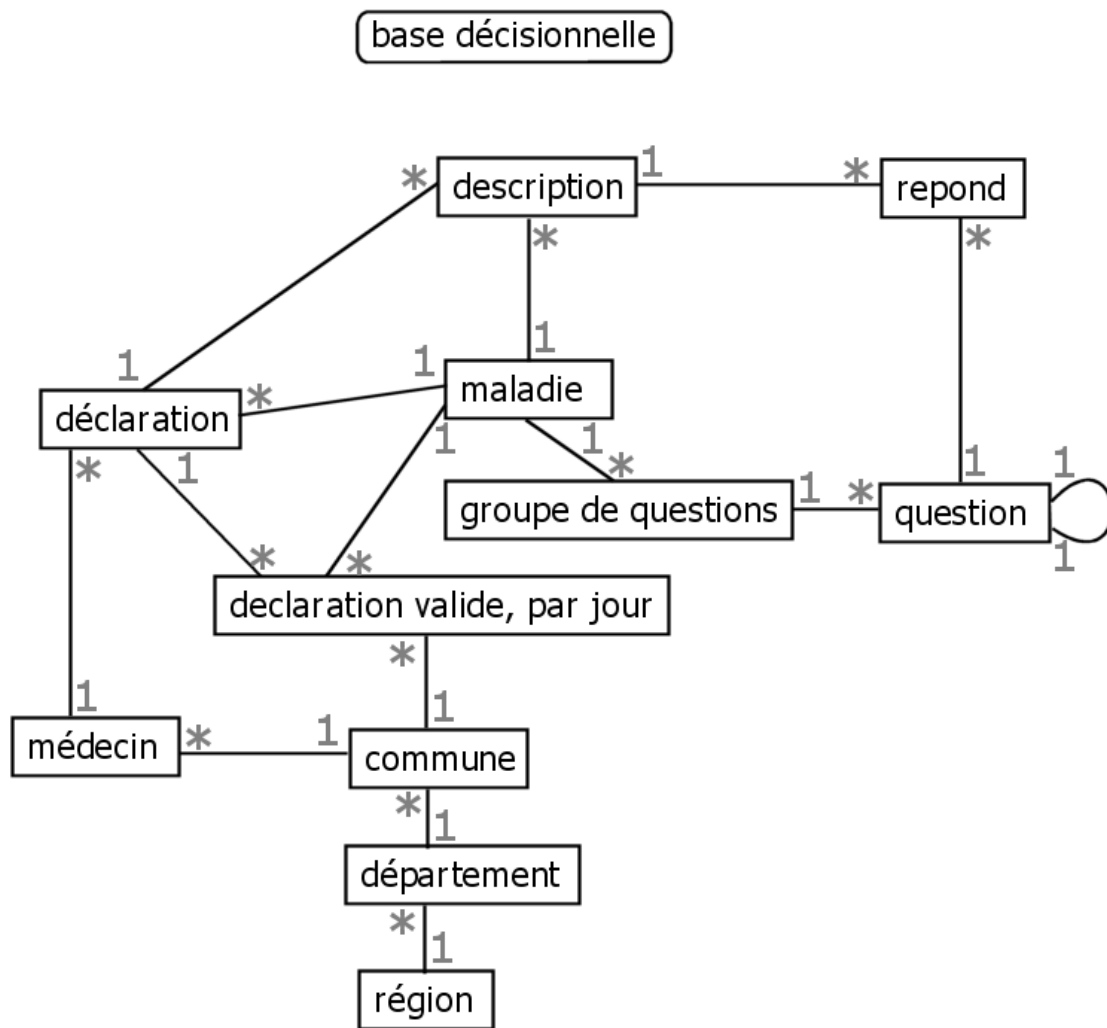


## Annexe 2 : Diagramme des fonctions du site Sentiweb

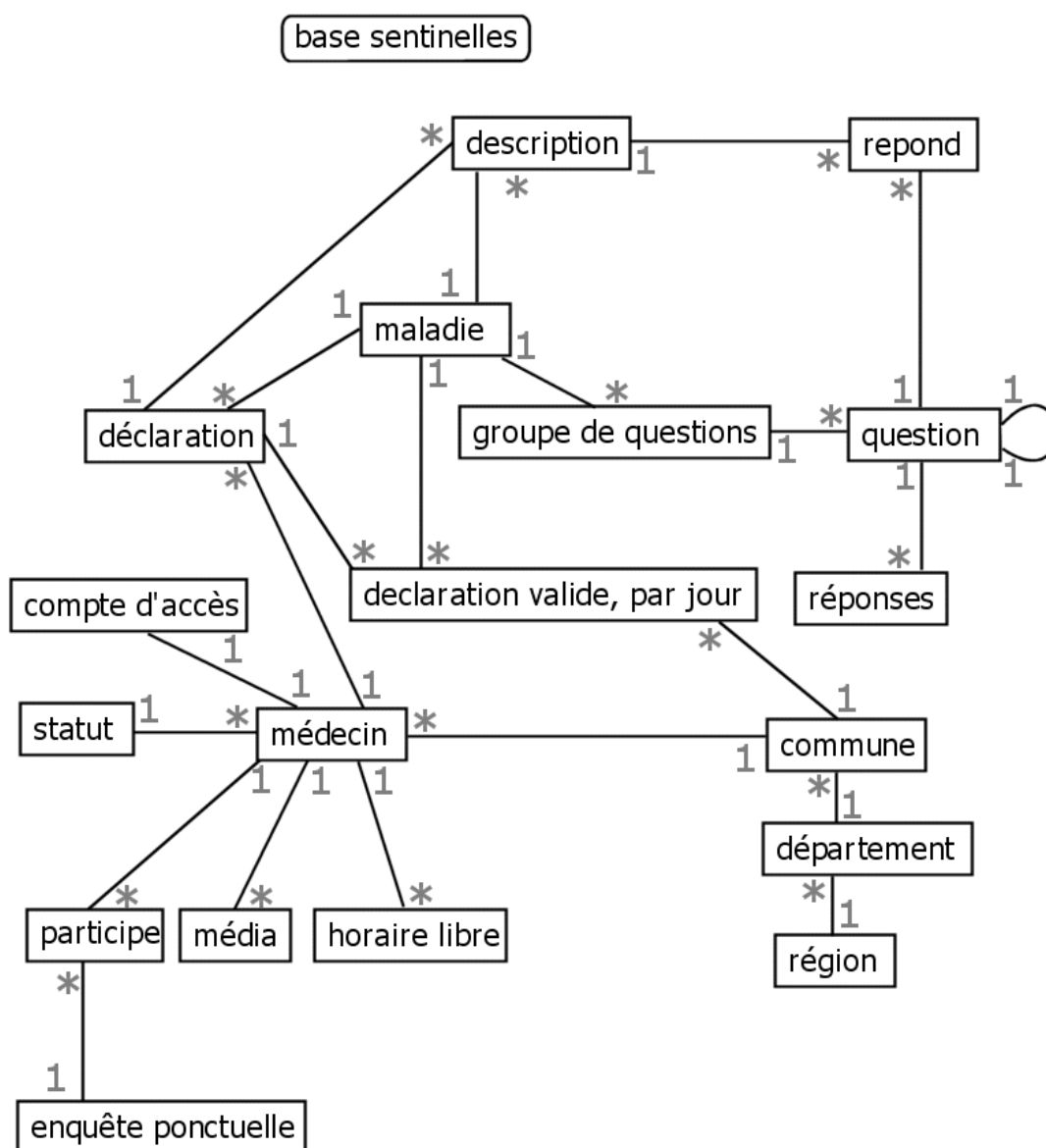




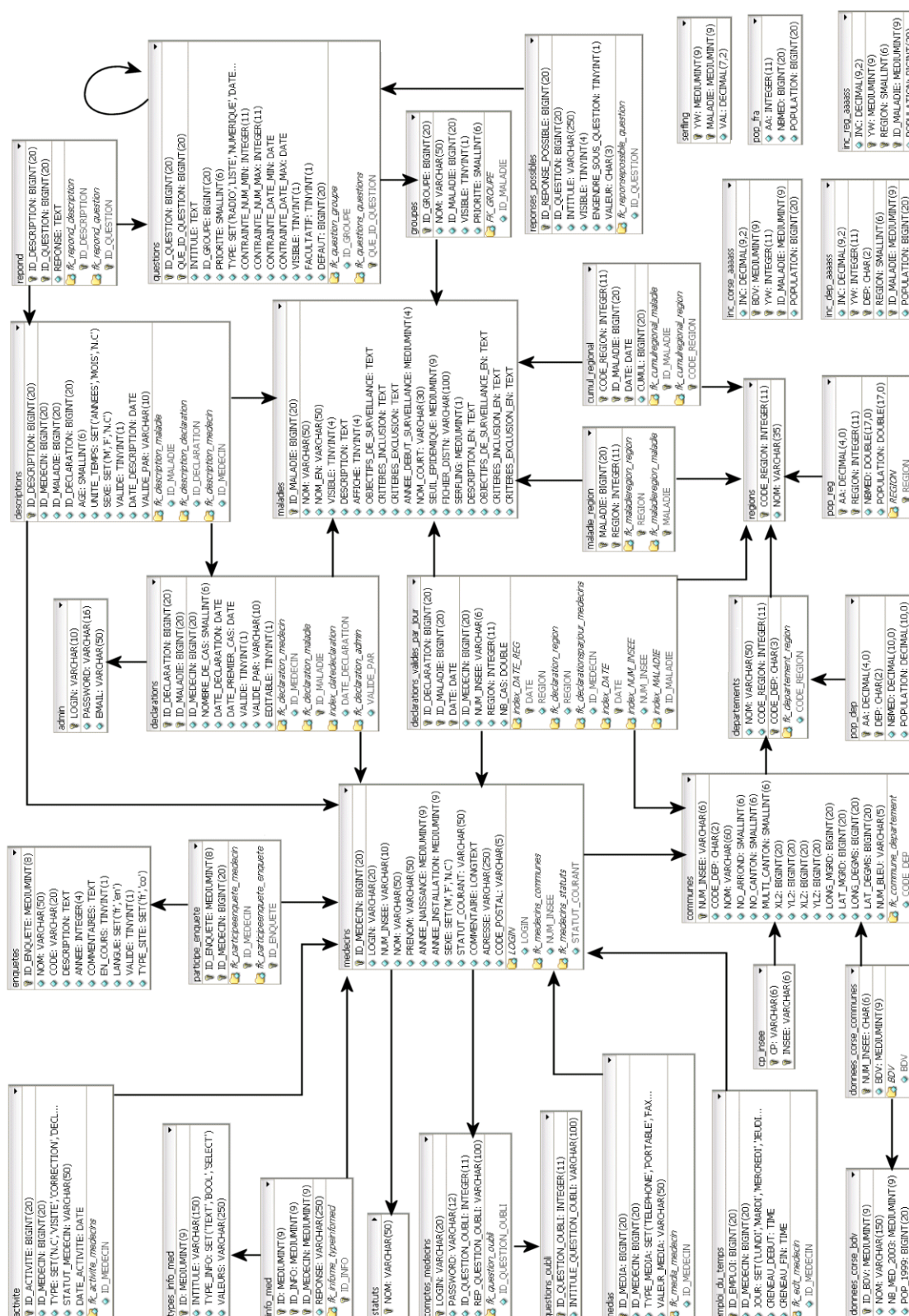
Annexe 4 : Diagramme UML simplifié de la BDD décisionnelle



Annexe 5 : Diagramme UML simplifié de la BDD Sentinelles unique



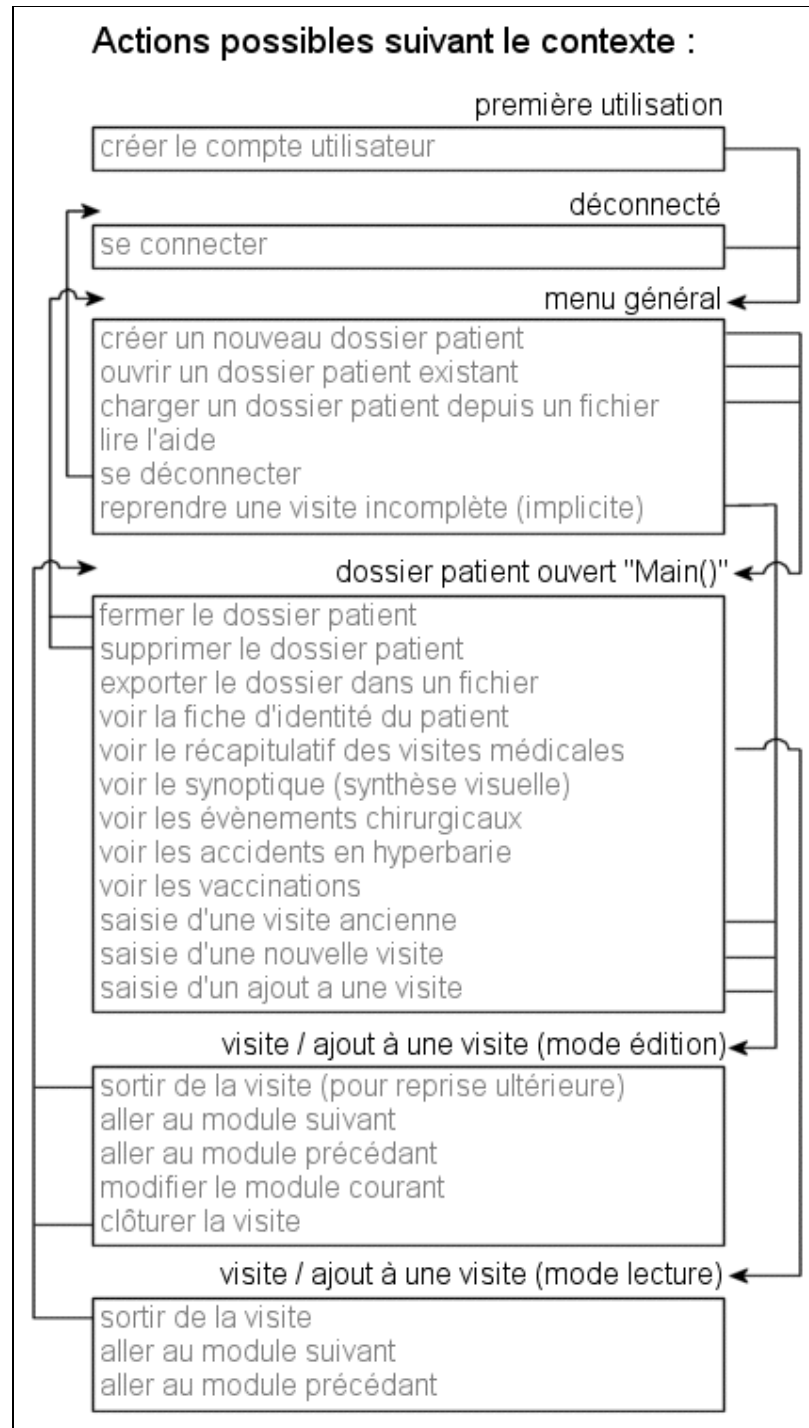
## Annexe 6 : Schéma complet de la BDD Sentinelles



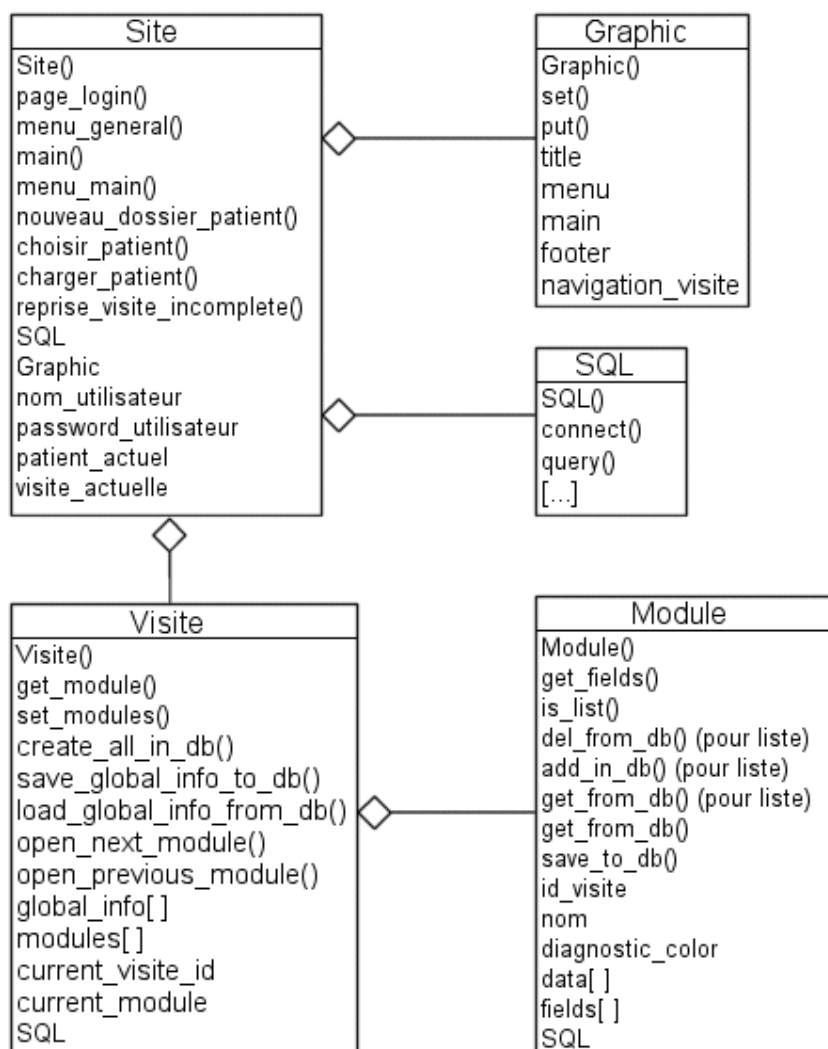


**GHB+**  
Gestion HyperBarie





## Annexe 9 : Schéma UML simplifié des classes de l'application GHB++



## Annexe 10 : Schéma UML de la BDD sous-jacente à l'application GHB++

