

Mise en place d'un entrepôt de données FHIR

16 juin 2024



Polytech Nantes

Filière *Ingénierie des données et de l'intelligence artificielle*

ITII Pays de la Loire



Xtrem Santé

4 rue Paul Vatine, 17180 Périgny
Groupe Softway Medical

Apprenant :

Erwan BRUNELLIÈRE - Promotion 2025
erwan.brunelliere@gmail.com

Référent Apprentissage :
Jean-Pierre ROYER

Tuteurs Industriel :

Yohann POIRON – Manageur Interopérabilité
Yohann.POIRON@esante.gouv.fr

Yannick LAGADEC – Engineering Director
ylagadec@softwaymedical.fr

Tutrice pédagogique :
Fatima-Zahra BOUCHANA

Table des matières

Mise en place d'un entrepôt de données FHIR.....	0
Table des matières	1
I. Présentation de l'entreprise.....	2
II. Raisons du projet	3
III. Présentation du projet et positionnement de l'apprenant.....	4
1. Analyse du standard	4
2. Analyse des solutions existantes.....	4
3. Modélisation et implémentation des outils.....	4
4. Développement d'un système de stockage et de recherche FHIR.....	5
5. Définition des objectifs et exigences de l'entrepôt	5
6. Construction de l'équipe.....	5
7. Evaluation de la conformité.....	5
9. Choix de la plateforme et de la technologie	6
11. Développement et test	6
12. Formation développeurs	6
13. Annonce Sante Expo	6
IV. Dimensions organisationnelles, humaines et économiques.....	7
1. Plan Organisationnel	7
2. Plan travail humain	7
3. Plan économique	7
V. Présentation du développement scientifique envisagé	8
Annexe 1.....	9
Annexe 2.....	10
Annexe 3.....	11

I. Présentation de l'entreprise

Xtrem Santé est le résultat d'un rachat (par le groupe Softway Medical) et d'une récente fusion.

L'entreprise est située à La Rochelle et à Chauray et comptabilise plus d'une centaine de collaborateurs.

Xtrem Santé a pour vocation de fournir aux cliniques et aux professionnels de médecine de ville un **système d'information** et de gestion des **données médicales** d'un parcours ville-hôpital.

Le domaine de l'entreprise est donc essentiellement l'**e-santé**. Elle propose 4 produits, adaptés aux professionnels de santé et quels que soient leurs lieux d'exercice : cabinets, centres et maisons de santé ou établissements privés.

Quant à elle, Softway Medical emploie près de 800 talents et équipe plus de 1000 établissements de santé. Elle est leader français de l'informatique en santé.

Ces deux entreprises permettent à leurs clients de manipuler **d'énormes quantités** de données médicales tous les jours.

La stratégie de Softway Medical est de **s'étendre** à l'international, en passant par le marché européen.

II. Raisons du projet

Ce projet arrive à la suite d'une avancée majeure dans le paysage des données de santé : l'arrivée du **standard FHIR** (Fast HealthCare Interoperability Resources). Elle a été mise en œuvre afin de répondre à plusieurs problématiques :

- **Volume et complexité des données de santé.** La quantité et la complexité des données de santé augmentent rapidement, provenant de diverses sources
- **Problème d'interopérabilité et de partage de données.** Le manque d'interopérabilité entre les systèmes de santé rend difficile le partage et l'utilisation efficaces des données, ce qui peut entraver la recherche, la prise de décision clinique et la prestation de soins de santé de qualité
- **Besoin d'un cadre standardisé pour l'échange de données.** Un cadre standardisé pour l'échange de données de santé est nécessaire pour faciliter la communication entre les systèmes disparates et promouvoir l'interopérabilité.

Le standard FHIR a pour objectif d'améliorer la manière dont les données médicales sont partagées et les rend plus facile d'accès et plus facile à interpréter. Sa structure est basée sur JSON, qui est autant **lisible** par un humain que par une machine.

Le format **normalisé** et la structure des ressources FHIR permettent un échange de données transparent entre les différents systèmes de santé.

FHIR favorise l'interopérabilité et le partage de données, éliminant les silos, ou monolithes de données.

La mise en place permettrait de fournir des données de santé, des données normalisées faciles à analyser.

Cela a énormément d'intérêt à l'heure actuelle, dans un monde préoccupé par le **Big Data**. Il fournirait un référentiel **centralisé** pour les données de santé provenant de **diverses** sources, permettant une **analyse** et des informations complètes.

Des données notamment utilisées par l'Intelligence Artificielle. On retient par exemple en cas d'usage :

- Les diagnostics assistés par IA
- Prédiction d'issue, de besoin de patients
- Optimisation des traitements
- Chats-bots pour conseils médicaux préliminaires
- Gestion de rendez-vous
- Prédiction des réadmissions d'un patient
- Détection précoce des maladies

Aujourd'hui le groupe Softway Medical n'exploite pas ces données médicales, elles ont pourtant une très grande valeur et il est une stratégie importante de pouvoir les exploiter.

Le groupe a donc pour stratégie de mettre en place son propre entrepôt de données afin d'exploiter ou de permettre l'exploitation des données qu'il possède.

III. Présentation du projet et positionnement de l'apprenant

Certaines tâches ont déjà été réalisées, d'autres sont à venir et sont dans une roadmap prévisionnelle, d'un point de vue XtremSanté et groupe Softway. Voir [Annexe 1](#), [Annexe 2](#) et [Annexe 3](#).

1. Analyse du standard

Le standard FHIR est complet et volumineux, il est important de le respecter. Chaque choix durant la création de l'entrepôt de données en lui-même devrait être guidé par la documentation (*Implementation Guide*) fournie.

Une bonne connaissance globale du standard ne peut qu'être appréciée lors du commencement de la modélisation et de l'implémentation.

2. Analyse des solutions existantes

Avant de se lancer dans la réalisation d'un projet, il est très important de se renseigner sur les solutions existantes. Certaines, comme l'*API Cloud Healthcare* de *Google*, sont payantes, d'autres gratuites et open source, le projet *HAPI FHIR* est une référence et est très utilisé pour les serveurs actuels. Nous avons pris en compte les solutions les plus utilisées, cherché à comprendre comment marchaient celles auxquelles nous avons accès.

Une recherche sur quelles technologies sont utilisées derrière ces entrepôts

Vaut-il mieux payer un service, payer une licence, héberger une solution open source ou finalement créer sa propre solution et l'héberger ?

3. Modélisation et implémentation des outils

La norme FHIR a besoin d'outils qui seront utiles quels que soient les projets que nous voudrions monter après. Nous avons donc réalisé un composant *vendor*, une sorte de **boîte à outils** que nous pourrions utiliser dans tous nos projets **PHP**, un langage informatique que nous utilisons presque exclusivement dans l'entreprise. Ce composant est **Open Source** et est disponible à tous dans la large bibliothèque en ligne <https://packagist.org/>.

Cette boîte à outils a été réfléchi par M. POIRON, mon tuteur, M. BERTONNIER, le lead dev de l'équipe et moi-même. Elle a ensuite été réalisée entièrement par moi-même sous la supervision de M. BERTONNIER.

Elle contient entre autres le **modèle** entier des ressources FHIR sous différentes versions, un outil pour les **lire** et les intégrer et un autre pour les **retranscrire**.

4. Développement d'un système de stockage et de recherche FHIR

Nous avons voulu essayer une première approche de l'entrepôt en l'incluant dans l'application maîtresse de l'entreprise.

J'ai pu **modéliser** et **implémenter** ce système de stockage, toujours suivi par la même équipe.

Nous avons pu tester ce système de stockage en reproduisant l'entrepôt de L'Annuaire de Santé **ANS**. Cette base sert à répertorier tous les professionnels de santé de France. Il est un entrepôt d'un point stratégique très important pour tous les éditeurs comme nous mais n'est à l'heure actuelle pas opérationnel pour être vraiment utilisé (il ne sert actuellement que de test).

5. Définition des objectifs et exigences de l'entrepôt

Identifier les **besoins spécifiques** en termes de **données**, les **objectifs** du projet, les **utilisateurs** finaux.

Définir le **type de données** à collecter (données cliniques, administratives, etc.), les **sources** potentielles et **l'usage** qu'il aura (données publiques, données privées, données publiques après anonymisation, etc.).

6. Construction de l'équipe

Etant donné qu'un tel entrepôt est un **axe stratégique** et que le groupe n'a, à part nous, pas la moindre piste de commencement, une équipe se constituera au sein du groupe, réunissant plusieurs filiales.

7. Evaluation de la conformité

Nous allons traiter des **données de santé**, nous devons alors faire très attention à ces données.

Nous devons prendre en compte le respect du **RGPD**. Le Règlement Général de Protection des Données est un **texte** qui **encadre** le traitement des **données** sur le territoire de l'Union Européenne.

Nous devons définir des protocoles de gestion des **données sensibles** et des consentements des patients.

Je serai amené à participer activement à cette tâche avec l'aide d'experts dans ce domaine. Cette partie du projet est très importante pour décider d'un cahier des charges et de la suite du projet.

C'est une partie qui n'est d'ailleurs pas exclusive à ce projet et que je serai sans aucun doute amené à expérimenter ultérieurement dans ma vie professionnelle.

8. Service de conversion de messages

De **nombreux standards** de la santé existent et l'arrivée de FHIR ne les fait pas disparaître. Beaucoup de systèmes d'information d'établissement de santé n'utilisent pas encore ce standard, mais il est nécessaire de pouvoir **communiquer** avec eux.

J'aurais la responsabilité de définir les stratégies possibles des technologies utilisées pour ensuite la mettre en place.

La bonne stratégie sera décidée par les acteurs décisifs du groupe.

9. Choix de la plateforme et de la technologie

Une société **externe** s'occupera d'analyser les différentes stratégies que nous aurons pu mettre en évidence. Notre solution maison, un service payant ou open source.

10. Sécurité et confidentialité des données

Nous devons implémenter un système pour gérer les autorisations.

Dans le cas d'un entrepôt public qui a le droit d'ajouter, de supprimer ou d'altérer des données. Dans le cas d'un entrepôt privé, gérer les **cloisonnements** et les **autorisations** de tous les utilisateurs.

Une tâche que je réaliserai avec l'aide de Florian Bertonnier.

11. Développement et test

Nous implémenterons l'entrepôt. Les efforts seront priorisés sur les **performances** de l'entrepôt.

Je serai bien évidemment amené à participer à cette tâche, sûrement accompagné d'une équipe.

12. Formation développeurs

Après toutes ces tâches réalisées autour du standard, nous aurons beaucoup de savoir autour de lui.

Nous **partagerons** notre savoir du standard FHIR aux autres développeurs pour qu'ils implémentent leurs prochaines solutions en utilisant cette norme ou en s'en rapprochant le plus possible.

13. Annonce Sante Expo

Le salon Santé Expo est un événement important pour le monde de la médecine. Je serai amené à **présenter** l'entrepôt et tous ses axes stratégiques qu'il apportera sur le stand de Softway Medical à côté de mes collaborateurs.

IV. Dimensions organisationnelles, humaines et économiques

1. Plan Organisationnel

Le projet sera amené à beaucoup changer en fonction des **objectifs** que l'entrepôt aura, en fonction des choix que le **groupe** aura vis-à-vis de cet entrepôt, quels rôles il aura, de quelles manières il sera utilisé et quelle stratégie sera employée pour sa conception.

Comment peut-on garder une cohérence d'équipe malgré un effectif réparti entre plusieurs filiales ?

2. Plan travail humain

La création d'un tel entrepôt impliquera plusieurs personnes.

Dans un premier temps la création et la gestion d'une **équipe** chargée de son implémentation. Avec des experts dans plusieurs domaines : des développeurs, ingénieurs des données, ingénieurs sécurité, etc.

Nous aurons plusieurs interactions avec d'autres acteurs extérieurs. Une analyse du **besoin** envers le groupe et de potentiels clients.

Une **formation** et **sensibilisation** des développeurs auprès du standard FHIR.

Une **présentation** de nos services, **convaincre** les clients d'utiliser notre service avec son avantage stratégique ou convaincre les utilisateurs de la qualité de nos données.

Comment se passera l'**adaptation** des chercheurs et des développeurs à cette nouvelle norme ?

3. Plan économique

L'entrepôt peut avoir une très grande ampleur, une très grande volumétrie de données. Il sera très utile, par la **valeur des données** qu'il contient.

Quel sera le prix de l'**accès** à cette bibliothèque de données ? Les **données médicales** (ici anonymisées) sont rares et il est important de savoir en tirer profit.

Il sera également important pour les **fonctionnalités** qu'il offrira. Ce sera un outil pour récupérer des données publiques, voire une référence.

Quel sera le prix **d'acquérir cet avantage** ?

Et au-delà d'un simple calcul de sommes, quelle sera la plus-value de l'entreprise à posséder un tel entrepôt ?

V. Présentation du développement scientifique envisagé

L'entreprise développe ses applications uniquement en utilisant le langage **PHP**. Or des entrepôts open source existent mais sont réalisés en Java, qui n'est pas du tout dans les compétences de notre entreprise. **Adapter** cet entrepôt selon nos besoins serait impossible sans une **formation** des développeurs.

Vaut-il mieux se **former** sur de nouvelles technologies uniquement dans le besoin de l'entrepôt ou vaut-il mieux développer **notre propre solution** ?

Si oui, en plus du PHP, quelles technologies utiliser, vaut-il mieux **greffer** ce projet sur notre application existante avec tous les outils déjà présents ou **créer** un projet à partir de presque rien pour ne pas être contraint et ralenti par d'autres outils ?

Une question se posera sur la base de données utilisée. **Laquelle prendre** ? De **quel type** ? Relationnelle ou NoSQL ?

En fonction de nos choix, choisir ensuite **l'architecture** de nos données, les **stratégies** utilisées.

Trouver un juste milieu entre performances, taille des données et persistance.

Un énorme travail **d'analyse** d'algorithmes, de stockage avec des tests de **performance** sur des infrastructures dédiées.

Une passe également sur l'anonymisation des données, pouvoir mettre à disposition de "tous" des données patient.

Un autre sujet important concerne la **sécurité** des données.

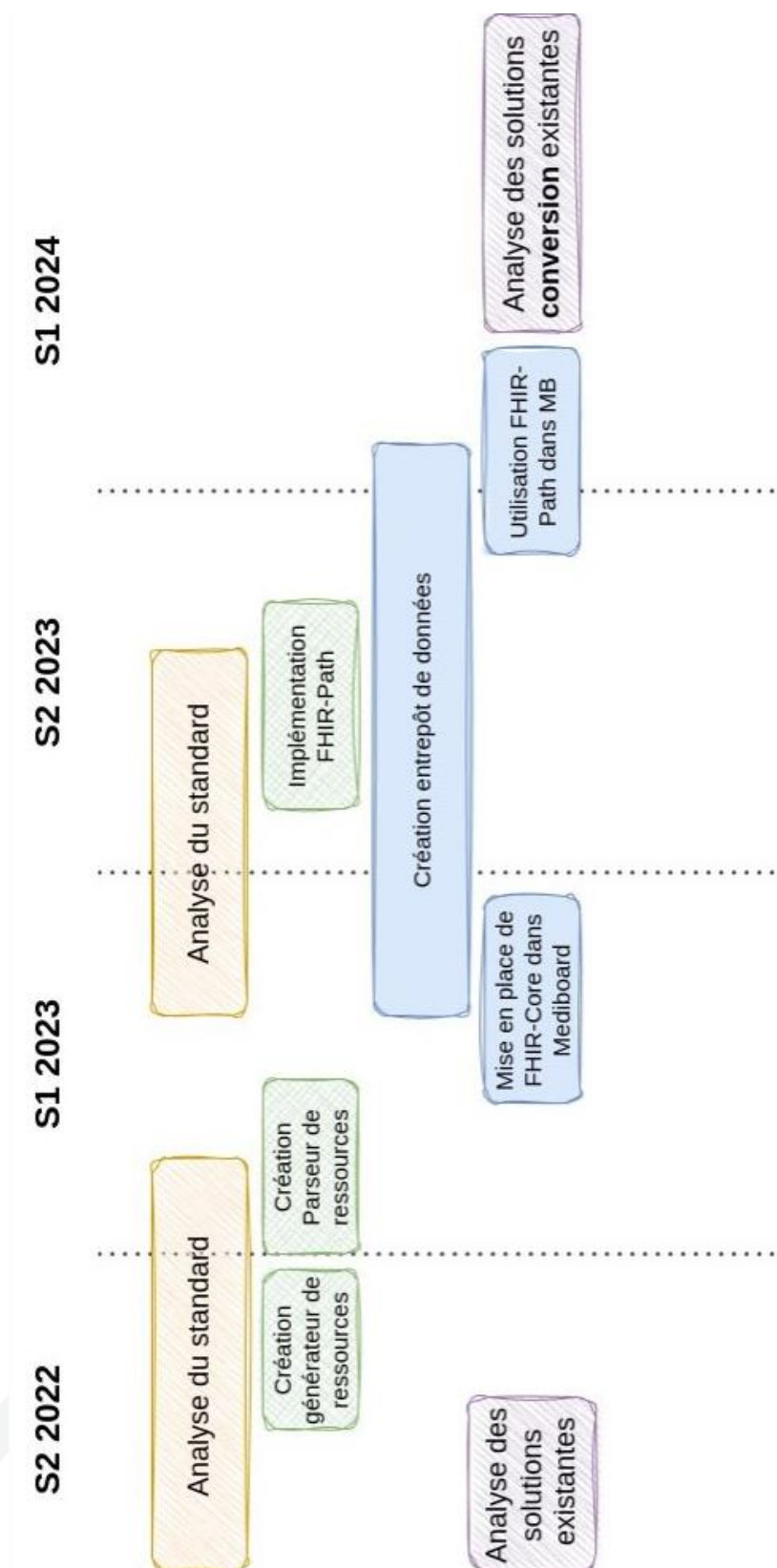
Garder nos données privées **confidentielles**, gérer les accès à l'entrepôt. Créer un système de **cloisonnement** des données pour séparer les utilisateurs au sein d'un même entrepôt.

Un système de **chiffrement** de données sera sans aucun doute également nécessaire.

Beaucoup de questionnement sur les différents **protocoles de sécurité, d'autorisation** et de **chiffrement**.

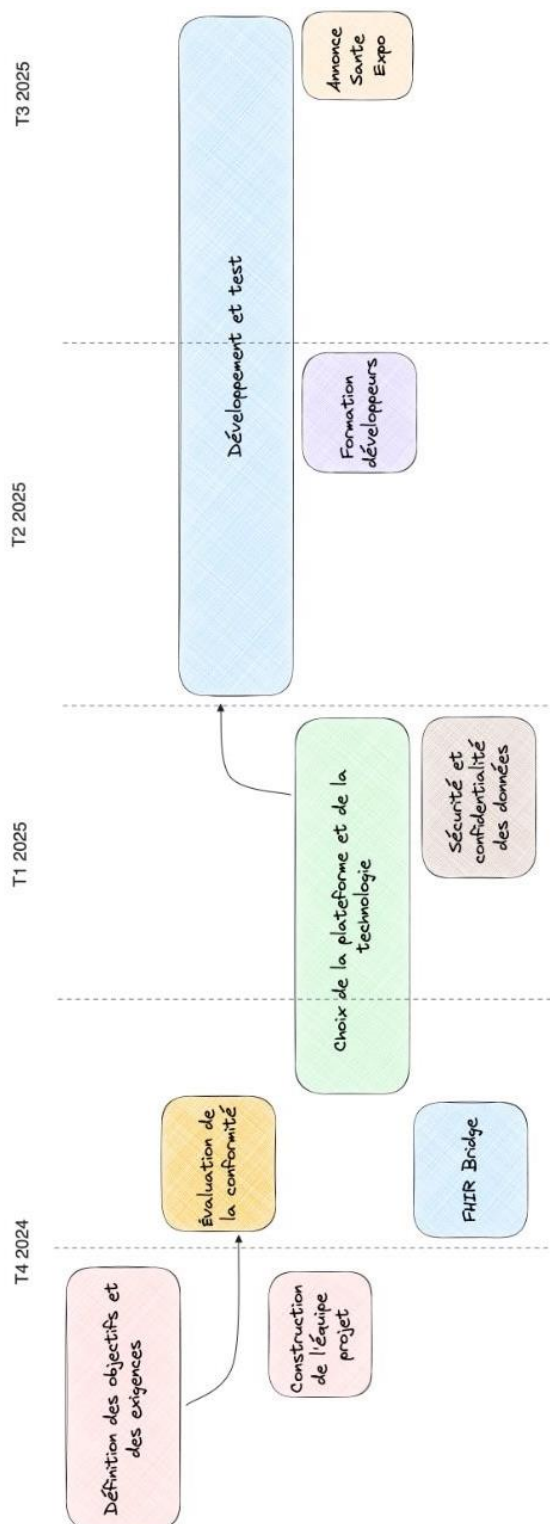
Annexe 1

Diagramme des tâches réalisées (Roadmap effective)



Annexe 2

Diagramme des tâches à réaliser (Roadmap cible) Groupe Softway



Annexe 3

Roadmap cible (non ordonnée) Xtrem Santé

