



République Algérienne Démocratique et Populaire  
Ministère de l'enseignement supérieur et de la  
recherche scientifique

Université Larbi Tébessi - Tébessa

Faculté des Sciences Exactes et des Sciences de la Nature et de la Vie  
Département : Mathématiques et Informatique



Mémoire de fin d'études  
Pour l'obtention du diplôme de **MASTER**  
Domaine : Mathématiques et Informatique  
Filière : Informatique  
Option : Systèmes et Multimédias

Thème

**Un modèle basé Blockchain pour un  
diagnostic médical en ligne efficace et  
préservant la confidentialité**

Présenté Par :  
Rais Khedidja

Devant le jury :

|                           |      |                          |              |
|---------------------------|------|--------------------------|--------------|
| Dr. Amroune Mohamed       | MCA  | Université Larbi Tébessi | Président    |
| Dr. Zeggari Ahmed         | MCB  | Université Larbi Tébessi | Examineur    |
| Pr. Derdour Makhlouf      | Prof | Université Larbi Tébessi | Encadreur    |
| Dr. Betouil Ali Abdelatif | MCB  | Université Larbi Tébessi | Co-Encadreur |

Date de soutenance : 29/06/2020

*À mes parents, à qui je dois trop.*

*Rais.K*

# Remerciement

---

---

*Ce travail n'a pu être mené à bien qu'avec le soutien de plusieurs personnes que je voudrais, à travers ces quelques lignes, remercier du fond du cœur.*

*Premièrement, nous remercions Dieu source de toute connaissance.*

*Je voudrais adresser toute ma gratitude au directeur de ce mémoire, Pr. Dourdour Makhlouf et Dr. Betouil Ali Abdelatif, enseignants au département de mathématiques et d'informatique, faculté des sciences naturelles et des sciences de la vie de l'université de Tébessa, pour la patience, la disponibilité et surtout les judicieux conseils, qui ont contribué à alimenter ma réflexion.*

*Je remercie Dr. Amroune Mohamed et Dr. Zeggari Ahmed, enseignants au département de mathématiques et d'informatique, faculté des sciences naturelles et des sciences de la vie de l'université de Tébessa, d'avoir accepté d'être les jury de cette thèse.*

*Enfin je désire aussi remercier tous les enseignants de l'Université Larbi Tebessi département de mathématiques et d'informatique, en me fournissant des données précises pour réussir à mes études universitaires.*

*Merci à tous.*

# Résumé

---

Les données personnelles des patients augmentent de façon exponentielle en raison de la numérisation des systèmes médicaux et du nombre croissant de patients et de maladies, qui peuvent être collectées lors du diagnostic en ligne, par exemple pour les utiliser à des fins commerciales dans la fabrication de médicaments ou pour la recherche scientifique sans leur autorisation, ainsi que d’envahir la vie privée du patient, connaissant son dossier médical, ce qui peut affecter sa psyché.

Un modèle de sécurité centralisé sera très difficile et coûteux à mettre à l’échelle, à entretenir et à gérer, ainsi qu’une infrastructure de sécurité centralisée introduira un point de défaillance unique et sera une cible facile pour les attaques DDoS (attaques par déni de service) et l’infrastructure centralisée sera difficile à mettre en œuvre dans une configuration industrielle où les nœuds périphériques sont largement répartis géographiquement.

Au cours de cette étude, nous proposons un nouveau système de conservation des données médicales basé sur la blockchain, nous démontrons la capacité de cette technologie à maintenir la confidentialité et la transparence des données stockées en ligne, avec des informations cryptées et protégées contre les accès et les utilisations indésirables et un système permettant de vérifier toutes les transactions et d’assurer la transparence de ces derniers.

Notre résultat à travers cette étude a été largement satisfaisant, car nous avons démontré la puissance de la technologie blockchain à deux niveaux de sécurité, à savoir l’intégrité et la confidentialité.

**Mots-clés :** données personnelles, diagnostic en ligne, sécurité, vie privée, collecte de données, blockchain, confidentialité , intégrité.

# Abstract

---

Patient personal data is growing exponentially due to the digitization of medical systems and the increasing number of patients and diseases which can be collected during online diagnostics for commercial use in the manufacture of medicines for example or for scientific research without their authorization as well as to invade the patient's privacy, knowing his medical record, which can affect his psyche.

A centralized security model will be very difficult and expensive to scale, maintain and manage because a centralized security infrastructure will introduce a single point of failure and be an easy target for DDoS attacks (Denial-of-service attacks) and centralized infrastructure will be difficult to implement in an industrial configuration where the peripheral nodes are widely distributed geographically.

During this study, we propose a new blockchain based medical data retention system, we demonstrate the ability of this technology to maintain the confidentiality and transparency of data stored online, with encrypted information protected against unwanted access and use, as well as a system for verifying all transactions and ensuring their transparency.

Our result through this study was largely satisfactory since we demonstrated the power of blockchain technology at two levels of security, namely integrity and confidentiality.

**Key words :** personal data, online diagnosis, security, privacy, data collection, blockchain, confidentiality, integrity.

## ملخص

تزداد البيانات الشخصية للمرضى بشكل كبير بسبب رقمنة الأنظمة الطبية والعدد المتزايد من المرضى والأمراض ، والتي يمكن جمعها أثناء التشخيص عبر الإنترنت ، على سبيل المثال للاستخدام التجاري في تصنيع الأدوية أو البحث العلمي دون إذن منهم ، وكذلك انتهاك خصوصية المريض بمعرفة سجله الطبي مما يمكن أن يؤثر على نفسيته.

سيكون نموذج الأمان المركزي صعبا للغاية ومكلفا من حيث الحجم، المحافظة عليه وإدارته ، بالإضافة إلى أن البنية التحتية الأمنية المركزية ستقدم نقطة العطل، كما أنها هدفا سهلا لهجمات الحرمان من الخدمات ، وسيكون من الصعب تنفيذ البنية التحتية المركزية في التكوين الصناعي حيث يتم توزيع العقيد على نطاق واسع جغرافيا.

خلال هذه الدراسة ، نقترح نظاما جديدا للاحتفاظ بالبيانات الطبية يستند على البلوكشين ، نثبت قدرة هذه التقنية في الحفاظ على سرية وشفافية البيانات المخزنة على الإنترنت ، مع معلومات مشفرة محمية ضد الوصول والاستخدام غير المرغوب فيهما ، بالإضافة إلى نظام التحقق من جميع المعاملات وضمان شفافيتها.

لقد كانت نتائجنا من خلال هذه الدراسة مرضية إلى حد كبير ، لأننا أظهرنا قوة تكنولوجيا البلوكشين على مستويين من الأمن ، وهما النزاهة والخصوصية.

**الكلمات المفتاحية:** البيانات الشخصية ، التشخيص عبر الإنترنت ، الأمان ، الحياة الشخصية ، جمع البيانات ، بلوكشين ، النزاهة ، الخصوصية.

# Table des matières

|  |             |
|--|-------------|
| <b>Remerciement</b>  | <b>i</b>    |
| <b>Résumé</b>  | <b>ii</b>   |
| <b>Table des matière</b>                                       | <b>v</b>    |
| <b>Table des figures</b>                                       | <b>viii</b> |
| <b>Liste des tableaux</b>                                      | <b>xi</b>   |
| <b>Introduction générale</b>                                   | <b>1</b>    |
| <b>1 Protection des données personnelles</b>                   | <b>3</b>    |
| 1.1 Introduction . . . . .                                     | 3           |
| 1.2 Contexte et historique . . . . .                           | 4           |
| 1.3 Données personnelles . . . . .                             | 5           |
| 1.3.1 Types des données personnels . . . . .                   | 6           |
| 1.3.2 Littératie des données . . . . .                         | 7           |
| 1.3.3 Processus de la littératie des données . . . . .         | 11          |
| 1.4 Classification des données personnelles . . . . .          | 11          |
| 1.5 Protection des données personnelles . . . . .              | 12          |
| 1.5.1 Raisons morales pour protéger les données personnelles   | 12          |
| 1.5.2 Loi, réglementation et contrôle indirect sur l'accès . . | 13          |
| 1.5.3 Données et contextes sensibles . . . . .                 | 14          |
| 1.5.4 La vulnérabilité . . . . .                               | 14          |
| 1.5.5 Violation des données . . . . .                          | 15          |
| 1.5.6 Raison d'attaque . . . . .                               | 17          |
| 1.5.7 Vie privée . . . . .                                     | 18          |
| 1.5.8 Confidentialité des données . . . . .                    | 20          |
| 1.6 Conclusion . . . . .                                       | 21          |
| <b>2 L'état de l'art</b>                                       | <b>22</b>   |
| 2.1 Introduction . . . . .                                     | 22          |
| 2.2 Sécurité . . . . .   | 22          |
| 2.2.1 Service de sécurité . . . . .                            | 23          |

|          |  |           |
|----------|--|-----------|
| 2.2.2    | Cryptographie . . . . .  | 24        |
| 2.2.3    | Techniques de cryptographie . . . . .  | 24        |
| 2.2.4    | Cryptanalyse . . . . .   | 27        |
| 2.2.5    | Fonction de hachage cryptographique . . . . .  | 27        |
| 2.3      | Blockchain . . . . .   | 29        |
| 2.3.1    | Historique . . . . .   | 29        |
| 2.3.2    | Composants de blockchain . . . . .   | 30        |
| 2.3.3    | Modèles de consensus . . . . .   | 36        |
| 2.3.4    | Contrats intelligents . . . . .  | 40        |
| 2.4      | Blockchain et la préservation des données . . . . .  | 42        |
| 2.4.1    | Système de conservation des données basé sur la blockchain pour les données médicales . . . . .  | 44        |
| 2.4.2    | DNS-IdM : un système de gestion d'identité Blockchain pour sécuriser le partage de données personnelles dans un réseau . . . . .                         | 46        |
| 2.4.3    | Une blockchain décentralisée de soins de santé préservant la confidentialité pour l'IoT . . . . .  | 47        |
| 2.4.4    | Système de Blockchain de Healthcare utilisant des contrats intelligents pour une surveillance à distance automatisée et sécurisée des patients . . . . . | 48        |
| 2.4.5    | Synthèse . . . . .   | 49        |
| 2.5      | Conclusion . . . . .   | 52        |
| <b>3</b> | <b>Contribution : une architecture basée blockchain et certificat intelligent pour protéger les données personnelles.</b>                                | <b>53</b> |
| 3.1      | Introduction . . . . .   | 53        |
| 3.2      | Système Healthcare . . . . .   | 53        |
| 3.3      | Conception de notre système . . . . .  | 54        |
| 3.3.1    | Diagramme d'activité . . . . .   | 55        |
| 3.3.2    | Diagramme d'états-transitions . . . . .  | 56        |
| 3.4      | Système cloud . . . . .  | 57        |
| 3.4.1    | Problème de confidentialité . . . . .  | 58        |
| 3.4.2    | Solutions suggérées . . . . .  | 59        |
| 3.4.3    | Niveaux de sécurité . . . . .  | 66        |
| 3.5      | Conclusion . . . . .   | 66        |
| <b>4</b> | <b>Implémentation et conception</b>  | <b>67</b> |
| 4.1      | Introduction . . . . .   | 67        |
| 4.2      | Scénarios d'étude . . . . .  | 67        |
| 4.2.1    | Diagramme de séquence de création d'un compte médecin traitant ou responsable de service . . . . .   | 67        |
| 4.2.2    | Diagramme de séquence de création d'un compte médecin . . . . .  | 68        |
| 4.2.3    | Diagramme de séquence de création d'un compte patient . . . . .  | 69        |



|                            |   |           |
|----------------------------|---|-----------|
| 4.2.4                      | Diagramme de séquence de prise de rendez-vous . . . . . | 70        |
| 4.2.5                      | Diagramme de séquence des consultations . . . . .       | 71        |
| 4.3                        | Outils de développement et langages . . . . .           | 74        |
| 4.4                        | Interfaces de notre plateforme . . . . .                | 76        |
| 4.4.1                      | L'administrateur . . . . .                              | 76        |
| 4.4.2                      | Médecins traitants . . . . .                            | 78        |
| 4.4.3                      | Responsables de services . . . . .                      | 83        |
| 4.4.4                      | Médecin . . . . .                                       | 86        |
| 4.4.5                      | Patient . . . . .                                       | 88        |
| 4.5                        | Stockage off-chain . . . . .                            | 91        |
| 4.6                        | Conclusion . . . . .                                    | 92        |
| <b>Conclusion générale</b> |   | <b>93</b> |

# Table des figures

|      |   |    |
|------|---|----|
| 1.1  | Lois et projets de loi nationaux complets sur la protection des données et la confidentialité 2018[8] . . . . . | 5  |
| 2.1  | Processus de cryptographie[26] . . . . .  | 24 |
| 2.2  | Techniques de cryptographie[23] . . . . .   | 25 |
| 2.3  | La cryptographie symétrique . . . . .   | 25 |
| 2.4  | La cryptographie asymétrique . . . . .  | 26 |
| 2.5  | Schéma fonctionnel de la fonction de hachage[21] . . . . .  | 28 |
| 2.6  | Exemple de transaction de crypto-monnaie[28] . . . . .  | 31 |
| 2.7  | Exemple de transaction des activités entre le médecin et le patient . . . . .                                   | 31 |
| 2.8  | Exemple de code QR[28] . . . . .  | 32 |
| 2.9  | Chaîne générique de blocs[28] . . . . .   | 36 |
| 2.10 | Exemple de la chaîne de texte «blockchain»[28] . . . . .  | 37 |
| 2.11 | Un cycle de vie d'un contrat intelligent se compose de quatre phases principales[29] . . . . .                  | 41 |
| 2.12 | L'opération de conservation des données[22] . . . . .   | 45 |
| 2.13 | L'architecture de système[10] . . . . .   | 46 |
| 2.14 | Réseau de superposition de ce système[14] . . . . .   | 47 |
| 2.15 | Conception de système[16] . . . . .   | 48 |
| 3.1  | Système Healthcare . . . . .  | 54 |
| 3.2  | Diagramme de contexte de notre système . . . . .  | 55 |
| 3.3  | Diagramme d'activité . . . . .  | 56 |
| 3.4  | Diagramme d'états-transitions . . . . .   | 57 |
| 3.5  | Exemple de la fragmentation mixte . . . . .   | 59 |
| 3.6  | La fragmentation des données selon le numéro de bloc . . . . .  | 60 |
| 3.7  | Bloc d'un médecin . . . . .   | 62 |
| 3.8  | Bloc d'un patient . . . . .   | 62 |
| 3.9  | Bloc d'un diagnostic . . . . .  | 63 |
| 3.10 | Exemple d'une chaîne de blocs médecin . . . . .   | 64 |
| 3.11 | Exemple d'une chaîne de blocs patient . . . . .   | 64 |
| 3.12 | Exemple d'une chaîne de blocs diagnostic . . . . .  | 65 |

---

|      |   |    |
|------|---|----|
| 4.1  | Diagramme de séquence de compte valide d'un médecin traitant / responsable de service . . . . .   | 68 |
| 4.2  | Diagramme de séquence de compte invalide d'un médecin traitant / responsable de service . . . . . | 68 |
| 4.3  | Diagramme de séquence de compte valide d'un médecin . . . . .                                     | 69 |
| 4.4  | Diagramme de séquence de compte invalide d'un médecin . . . . .                                   | 69 |
| 4.5  | Diagramme de séquence de compte valide d'un patient . . . . .                                     | 70 |
| 4.6  | Diagramme de séquence de compte invalide d'un patient . . . . .                                   | 70 |
| 4.7  | Diagramme de séquence de prise de rendez-vous valide . . . . .                                    | 71 |
| 4.8  | Diagramme de séquence d'un rendez-vous non valide . . . . .                                       | 71 |
| 4.9  | Diagramme de séquence d'une consultation valide . . . . .   | 72 |
| 4.10 | Diagramme de séquence d'une consultation invalide . . . . .                                       | 72 |
| 4.11 | Exemple 1 d'une consultation invalidée . . . . .  | 73 |
| 4.12 | Exemple 2 d'une consultation invalidée . . . . .  | 73 |
| 4.13 | Exemple 3 d'une consultation invalidée . . . . .  | 73 |
| 4.14 | Exemple 4 d'une consultation invalidée . . . . .  | 74 |
| 4.15 | Nouveaux responsables de services . . . . .   | 76 |
| 4.16 | Nouveaux médecins traitants . . . . .   | 77 |
| 4.17 | Responsables de services validés . . . . .  | 77 |
| 4.18 | Médecins traitants validés . . . . .  | 78 |
| 4.19 | Exemple d'un compte médecin traitant avant la validation . . . . .                                | 78 |
| 4.20 | Exemple d'un compte médecin traitant refusé . . . . .   | 79 |
| 4.21 | Profil d'un médecin traitant . . . . .  | 79 |
| 4.22 | Nouveaux patients . . . . .   | 80 |
| 4.23 | Patients validés . . . . .  | 80 |
| 4.24 | Nouveaux diagnostics . . . . .  | 81 |
| 4.25 | Exemple d'un revue . . . . .  | 81 |
| 4.26 | Diagnostics validés . . . . .   | 82 |
| 4.27 | Informations de diagnostic . . . . .  | 82 |
| 4.28 | Informations de patient . . . . .   | 82 |
| 4.29 | Informations des médecin . . . . .  | 83 |
| 4.30 | Profil d'un responsable de service . . . . .  | 83 |
| 4.31 | Nouveaux médecins . . . . .   | 84 |
| 4.32 | Médecins validés . . . . .  | 84 |
| 4.33 | Nouveaux diagnostics . . . . .  | 85 |
| 4.34 | Diagnostics validés . . . . .   | 85 |
| 4.35 | Gestion des rendez-vous . . . . .   | 86 |
| 4.36 | Profil d'un médecin . . . . .   | 86 |
| 4.37 | Création de diagnostic . . . . .  | 87 |
| 4.38 | Diagnostics validés . . . . .   | 87 |
| 4.39 | Diagnostics non valide . . . . .  | 88 |
| 4.40 | Rendez-vous par jour . . . . .  | 88 |
| 4.41 | Profil d'un patient . . . . .   | 89 |
| 4.42 | Prise de rendez-vous . . . . .  | 89 |

|      |  |    |
|------|--|----|
| 4.43 | Diagnostics validés . . . . .                | 90 |
| 4.44 | Rendez-vous . . . . .                        | 90 |
| 4.45 | Remboursement . . . . .                      | 91 |
| 4.46 | Fragmetation de données . . . . .            | 92 |
| 4.47 | Exemple de duplication des données . . . . . | 92 |

# Liste des tableaux

|     |  |    |
|-----|--|----|
| 2.1 | Comparaison des paramètres SHA[21]                             | 29 |
| 2.2 | Valeurs de digest SHA-256 correspondantes à le texte entré[28] | 30 |
| 2.3 | Comparaison des systèmes                                       | 51 |
| 4.1 | Outils de développement et langages de programmation           | 76 |

# Introduction générale

---

---

La technologie est devenue une partie intégrante des soins de santé parce qu'elle joue un rôle important dans le développement et la facilitation de ses services. La plupart des hôpitaux et des secteurs de la santé utilisent des systèmes médicaux numérisés au lieu des systèmes classiques. néanmoins, cette numérisation a créé d'autres problèmes :

- L'utilisation de bases de données locales pour stocker des informations médicales des patients n'est pas efficace et risque d'être accessible par des pirates.
- Avec la perspective d'une croissance exponentielle des structures de santé et des maladies ainsi que le nombre des malades, il est très difficile d'identifier, d'authentifier et de sécuriser les données personnelles des patients et probablement la vie privée des personnes.
  - Un modèle de sécurité centralisé sera très difficile et coûteux à mettre à l'échelle, à entretenir et à gérer.
  - Une infrastructure de sécurité centralisée introduira un point de défaillance unique et sera une cible facile pour les attaques DDoS.
  - Une infrastructure centralisée sera difficile à mettre en œuvre dans une configuration industrielle où les nœuds périphériques sont largement répartis géographiquement.
- Le remplacement des bases de données locales par des bases de données cloud pour pouvoir stocker plus de données présente un grand risque et un problème de confiance par rapport à la délocalisation des données dans des data-center éloignés.

Pour l'utilisation des systèmes de santé en ligne, la technologie de la blockchain semble être une alternative viable en raison des points forts décrits ci-dessus :

- 1 Elle peut être utilisée pour créer un réseau maillé sécurisé qui permettra aux médecins et malades de se connecter de manière sécurisée et fiable en évitant les menaces d'usurpation et d'emprunt d'identité.
- 2 Chaque consultation peut être enregistrée dans la blockchain et aura un identifiant de blockchain qui identifiera de manière unique le patient et le médecin dans l'espace de noms universel. Pour qu'un patient passe une autre consultation, on utilisera l'ID de la blockchain comme URL et utilisera son portefeuille de blockchain local pour soulever une

demande de consultation. Le portefeuille créera une demande signée numériquement et enverra au responsable de service qui utilisera les services de la chaîne de blocs pour valider le RDV à l'aide de la clé publique de l'expéditeur. De cette façon, l'authentification P2M (patient au médecin) peut avoir lieu sans avoir besoin d'un arbitre ou d'un service centralisé.

- 3 La solution possible ci-dessus sera applicable à une large gamme de services de santé. Certains des exemples seront les véhicules connectés aux Healthcare intelligents, la logistique, la télésurveillance, le diagnostic etc.

Dans cette thèse, nous proposons un nouveau système efficace basé sur la Blockchain pour préserver les diagnostics en ligne, intégrant d'autres approches pour améliorer le niveau de sécurité et les transactions :

- Un nouveau modèle de bloc différent du bloc standard, où les blocs sont liés les uns aux autres de manière à former une nouvelle chaîne de blocs.
- Les contrats intelligents qui nous permettent de contrôler les processus de diagnostic et les remboursements en temps réel éliminant la nécessité de passer par un intermédiaire.
- Nous utilisons la fragmentation des données dans un système de base de données distribué, pour désorganiser les données et réduire les risques de sécurité dans le cloud.

Le premier chapitre fournit un aperçu de la protection des données personnelles, explique les données personnelles, leur classification et tout ce qui sert à leur protection.

Le deuxième chapitre parle de la sécurité, la technologie blockchain, certains travaux qui ont utilisé la technologie blockchain pour la protection des données et une étude de synthèse pour clarifier l'état de l'art.

Le troisième chapitre présente l'essentiel de la contribution de ce travail, à savoir : une architecture basée blockchain pour préserver la vie privée des patients et un nouveau modèle de bloc pour renforcer le niveau de sécurité et de confidentialité des données privées.

Le quatrième chapitre présente l'implémentation et montre les différents scénarios de notre système.

# Chapitre 1

## Protection des données personnelles

---

*"Nous avons la responsabilité de protéger vos données et si nous ne pouvons pas le faire, nous ne méritons pas de vous servir."*

*Mark Zuckerberg*

---

### 1.1 Introduction

Avec les avancées technologiques et la numérisation des services publics et de santé, la protection des données est devenue d'une grande importance, d'autant plus que les données personnelles sont devenues faciles à collecter et à traiter sans l'autorisation de leur propriétaire et sa connaissance du traitement de ses données ou de l'endroit où elles peuvent être utilisées, ainsi que la possibilité d'accéder et de visualiser ses données ce qui touche à sa vie privée. Il est donc très important de prévenir les violations de données, de limiter leur utilisation et de maintenir la confidentialité.

Dans ce chapitre, nous présentons un contexte et historique sur la protection des données, nous définissons les données personnelles et leurs types en fonction de la littératie des données que nous allons expliquer leurs approches et processus, ainsi que nous donnons une classification des données personnelles, et enfin nous traitons la protection des données selon les raisons morales et juridiques, données et contextes sensibles, vulnérabilité des données, violation des données, raisons d'attaque, la vie privée et la confidentialité des données.



## 1.2 Contexte et historique

Les conversations théoriques et juridiques sur la relation entre la technologie et la vie privée remontent aux années 1890 avec l'avènement de l'équipement de photographie portable accessible à la population en général. Alors que les technologies continuent de se développer, des conceptualisations de la vie privée se sont développées parallèlement à elles, allant du "droit d'être laissé seul" aux tentatives de saisir la complexité des questions de vie privée dans des cadres qui mettent en évidence les préoccupations juridiques, socio-psychologiques, économiques ou politiques que les technologies présentent[15].

En 1948, la déclaration universelle des droits de l'homme (DUDH) a reconnu la vie privée comme un droit humain international pour la première fois, déclarant que "nul ne sera l'objet d'immixtions arbitraires dans sa vie privée, sa famille, son domicile ou sa correspondance, ni des atteintes à son honneur et à sa réputation." [17], et en 1974, la loi américaine sur la protection des renseignements personnels maintient les restrictions sur les données détenues par le gouvernement[25].

Dès 1988, le Comité des droits de l'homme des Nations Unies, l'organe conventionnel chargé de surveiller la mise en œuvre du pacte international relatif aux droits politiques et civils (PIDCP), a reconnu la nécessité de lois sur la protection des données pour sauvegarder le droit fondamental à la vie privée reconnu par l'article 17 de la le PIDCP[8].

Au milieu des années 1990, Roger Clarke a été le premier spécialiste de la protection de la vie privée à classer ses types de manière logique, structurée et cohérente[15].

La première protection des informations sur la santé remonte à 1996 (HIPAA) Health Insurance Portability and Accountability Act. En 1999 , GLBA qui signifie Gramm-Leach-Bliley Act a été établie dans bute de protéger les informations financières personnelles non-publiques (NPI). Le COPPA a été créé en 2000 pour protéger les données des enfants ( $\leq 12$  ans) ainsi que les règles de la confidentialité pour renforcer la HIPAA et protéger les informations de santé privées des individus. La loi SOX (Sarbanes-Oxley Act) qui protège le publique des pratiques fraudulentes des corporations ainsi que la loi FISMA qui représente l'acte de gestion des informations de sécurité fédérales en ordonnant les agences de protéger les données ont été formées en 2002. ISO 27001 fonctionne comme un cadre pour un système de gestion de la sécurité de l'information créé en 2013. En 2018 , La GPRD (la régulation générale de confidentialité des données) a été formée afin de protéger les données personnels des citoyens européennes[25] ainsi que plus de 100 pays à travers le monde avaient promulgué une législation complète sur la protection des données, et environ 40 pays étaient en train de promulguer de telles lois[8].

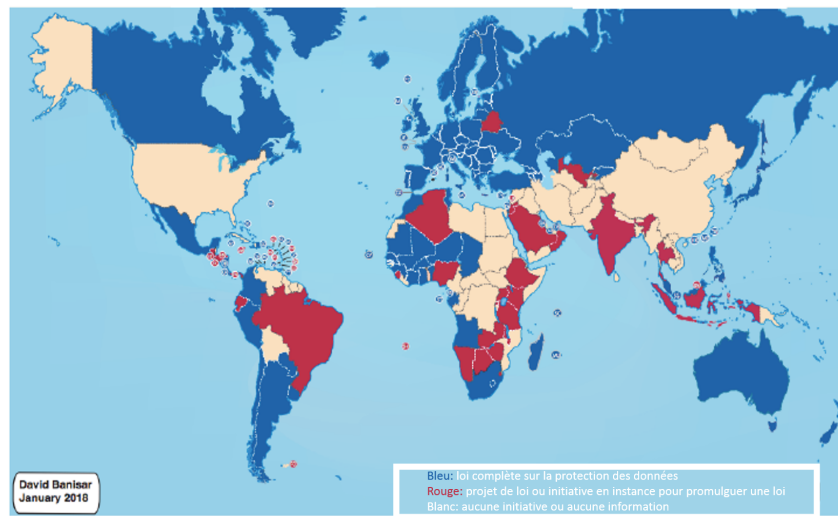


FIGURE 1.1: Lois et projets de loi nationaux complets sur la protection des données et la confidentialité 2018[8]

La loi sur la protection des informations personnels des consommateurs en californie a été établie en 2020 afin de restreindre la façon dont les companies collectent et utilisent les données[25].

### 1.3 Données personnelles

Les données personnelles sont tout élément d'information qui peut identifier ou être identifiable pour un individu. Celles-ci sont souvent appelées en termes juridiques "informations personnellement identifiables". Les formes numériques de données personnelles qui peuvent être extraites d'un large éventail de sources logicielles et matérielles revêtent une variété de modes, notamment les chiffres, les caractères, les symboles, les images, les ondes électromagnétiques, les informations des capteurs et les sons, présentent un intérêt particulier. Les données personnelles sont générées et rassemblées pour un large éventail de raisons, de l'amélioration des performances individuelles à des fins de sûreté et de sécurité. En particulier, une économie florissante de la connaissance des données s'est développée dans laquelle les données, souvent de nature personnelle, ont acquis une valeur commerciale considérable. Cependant, la définition des données personnelles n'est en aucun cas simple, «les données personnelles doivent être comprises comme une classe d'informations beaucoup plus vaste et encore plus envahissante que les éléments simples que nous pourrions penser».[24].

### 1.3.1 Types des données personnels

En développant un cadre pour la littératie des données personnelles, "capacité d'identifier, collecter, traiter, analyser et interpréter les données", il existe au moins trois types distincts de données personnelles[24].

**A. Données que les utilisateurs donnent aux appareils et systèmes :**

cela peut inclure des informations d'auto-suivi, des données de médias sociaux (y compris des vidéos, des images, des textes et des tweets), des e-mails et des vidéos. La prolifération des plateformes de médias sociaux telles que Facebook, Instagram et Twitter a augmenté le potentiel de collecte de données personnelles que les individus transmettent consciemment aux appareils et systèmes. Les pratiques adoptées par le biais de ces plateformes sont souvent expressives et émotives, ce qui donne lieu à des sources d'informations personnelles riches et détaillées. Les données personnelles peuvent également être générées volontairement par le biais d'activités qui se déroulent dans le cadre professionnel et éducatif. L'utilisation d'un "système de gestion de l'apprentissage" à l'université, par exemple, implique que les étudiants accèdent et téléchargent des ressources pédagogiques, contribuent à des forums de discussion et répondent à des questionnaires et à des tâches d'évaluation. Bien que ces activités puissent faciliter l'apprentissage, elles génèrent également des données personnelles sur chaque individu, qui peuvent être traitées et analysées pour prédire et optimiser l'utilisation future du système[24].

**B. Données que les appareils et systèmes extraient des utilisateurs :** sont involontaires et comprennent la "surveillance et la récolte de l'utilisation des appareils des personnes, les recherches et les transactions en ligne par les services de police et de sécurité, les empires Internet et l'industrie de l'exploration de données, et le développement d'outils et de logiciels pour produire, analyser, représenter et stocker de gros ensembles de données". Surtout, ces données sont créées grâce à leur collecte, ce qui signifie que l'entreprise, l'organisation ou l'institution qui génère ces données personnelles peut revendiquer le contrôle. L'individu dont les actions déclenchent la génération de ces données personnelles extraites a souvent le moins de contrôle, car les entreprises, les gouvernements, les chercheurs et les scientifiques cherchent à traiter et à réutiliser les informations collectées. L'interface de programmation d'applications (API) des plateformes numériques garantit que les entreprises technologiques ont le plus grand accès à ces données personnelles. Cela laisse les utilisateurs individuels toujours en déficit, jouant avec les spécialistes des données. Si une personne a accepté les "termes et conditions" d'une plateforme, cela est considéré comme une utilisation légale de ses données personnelles. Cependant, la question de savoir si les accords sur les termes et conditions peuvent

couvrir toutes les manières possibles de réutiliser les données[24].

- C. Données que les appareils et systèmes traitent pour le compte des utilisateurs :** sont les nombreuses façons dont les données personnelles sont traitées sous la forme d'entités de données plus significatives sur le plan social qui fournit des informations pertinentes pour les personnes, les lieux et les institutions. Souvent, les utilisateurs individuels seront peu ou pas exposés à ces données, car elles sont utilisées pour informer les processus du système et les procédures institutionnelles, souvent sous la forme de "doubles de données" conçues pour rendre l'individu identifiable, connaissable mais aussi interpassif. Pourtant, certaines formes de ces données sont renvoyées aux individus, bien que sous une forme partielle. Par exemple, de nombreuses plateformes numériques sociales et / ou grand public sont conçues pour agréger et traiter les données générées par la participation à la plateforme et les (re) présenter aux utilisateurs individuels via des tableaux de bord, des pages d'analyse et similaires. Les individus peuvent ensuite utiliser ces informations pour participer différemment dans le contexte de la plateforme. Malgré les efforts déployés pour présenter les données d'une manière facilement compréhensible par les utilisateurs (c'est-à-dire les visualisations, les tableaux de bord et les profils), l'identification et l'interprétation des données traitées peuvent être difficiles. Cela reflète en partie le fait que toute instance de traitement de données est dirigée par les motivations de ceux qui détiennent et contrôlent les données façonnant ainsi toute interprétation et utilité probables[24].

### 1.3.2 Littératie des données

Les approches actuelles de la littératie des données numériques "capacité d'identifier, collecter, traiter, analyser et interpréter les données", peuvent être considérées selon quatre formes.

**A. Sécurité et gestion des données :**

une approche de plus en plus populaire de la maîtrise des données est celle de la "sécurité des données" ancrée dans l'idée de protéger et / ou d'empêcher de manière sélective la diffusion des données personnelles. Cela étend les approches de "cybersécurité" et de "sécurité Internet" en se concentrant sur le développement des compétences des personnes pour gérer et contrôler les traces de données et les traces qu'elles laissent lors de l'utilisation des médias numériques, ce que l'on appelle parfois "l'empreinte numérique". Les approches de sécurité des données ont tendance à se concentrer sur les données personnelles que les individus fournissent volontairement aux appareils et systèmes. Cela peut être considéré comme des données sociales générées en utilisant les fonctionnalités des plateformes multimédias connectives, telles

que la création d'un profil personnel ou l'utilisation de boutons sociaux pour "J'aime", "voter positivement", évaluer, etc. La sécurité des données s'adresse le plus souvent aux enfants, aux jeunes et à leurs familles, en adoptant une approche didactique qui vise à améliorer les pratiques de vie privée et de sécurité. Comme pour la cybersécurité, le discours sur la "sécurité des données" a tendance à être promu le plus vigoureusement par les groupes de défense et les établissements d'enseignement, souvent sous la forme de sites Web et de programmes en classe. Celles-ci prennent généralement en charge le développement de stratégies normatives pouvant être mises en œuvre au sein d'un appareil ou d'un système, telles que l'ajustement des paramètres de sécurité au sein de la plate-forme ou la lecture des termes et conditions attachés à une plateforme particulière.

Une autre approche populaire de sécurité des données consiste à utiliser un logiciel qui permet à l'utilisateur de choisir avec quelles entreprises et marques il partage ses données personnelles. Plutôt que de modifier les paramètres de sécurité au sein des plateformes, des applications comme Citizenme et People.io visent à fournir aux utilisateurs une certaine forme de "contrôle" sur toutes les données personnelles qu'ils génèrent. Ces logiciels visent à "libérer" les données personnelles en mettant à la disposition des individus l'intelligence artificielle qui peut être collectée à partir de ces informations. Les utilisateurs bénéficient ainsi d'un aperçu personnel de leurs identités et interactions numériques. De plus, les utilisateurs peuvent être rémunérés (c'est-à-dire via PayPal, iTunes) pour partager leurs données avec les entreprises et les marques de leur choix[24].

- B. **Science des données** : ailleurs, c'est un soutien pour les formes alphabétisation basée sur la lecture, la compréhension et l'analyse d'ensembles de données ouvertes. Cette forme d'alphabétisation des données est axée sur le soutien de la capacité des individus à travailler avec de grands ensembles de données "ouverts" qui ont été collectés par les gouvernements et les organisations à diverses fins. Des exemples de l'approche de la science des données peuvent être trouvés dans le Data Boot Camp de la Banque mondiale, le Software Carpentry's Data Carpentry et le School of Data's Data Expeditions. Celles-ci sont toutes basées sur des approches pédagogiques "pratiques" allant des compétences de base en interprétation à un engagement complexe avec l'analyse de l'information géographique (IG) et des introductions structurées aux principes du développement de logiciels. Bien que la compréhension et l'analyse de ces données soient communément associées à la "maîtrise des données", il s'agit plus précisément d'une forme de "science des données" dans la mesure où elles impliquent des "statistiques, ou l'étude systématique de l'organisation, des propriétés

et de l'analyse des données et de leur rôle. en déduction".

Ces approches de la science des données à la maîtrise des données s'adressent généralement aux professionnels (tels que les journalistes, la société civile et les codeurs civiques) qui ont peu ou pas d'expérience préalable en programmation informatique. Par exemple, le programme Data Boot Camp de la Banque mondiale fait partie d'un programme plus large visant à "inspirer et habiliter les citoyens à utiliser les données ouvertes et à maximiser la valeur pour le public de manière pratique". En tant que telles, ces approches de la science des données à la maîtrise des données sont souvent étayées par un discours sur la productivité, centré sur le développement de compétences pour améliorer la société et soi-disant autonomiser les citoyens. Ces résultats semblent provenir des types de données avec lesquelles la science des données travaille le plus souvent, c'est-à-dire de grands ensembles de (méta) données collectées par les gouvernements et les organisations, et donc pertinentes pour les grands problèmes de société. Malgré la rhétorique de la transparence promise par la science des données "ouvertes", certains chercheurs soutiennent qu'il existe des contradictions inhérentes à la notion de gouvernements et d'autres intérêts acquies soutenant des formes significatives d'ouverture, d'accès libre et de données ouvertes[24].

- C. **Piratage de données** : un troisième type de la littérature des données est ce que l'on peut qualifier d'approches de "piratage des données" pour accéder aux données et les réaffecter. Alors que le piratage est souvent communément compris comme impliquant une introduction illégale dans les systèmes gouvernementaux ou d'entreprise, la philosophie du piratage au sein des cultures des programmeurs est plus abstraite dans ses objectifs. En ce sens, la philosophie du pirate est plus précisément perçue comme s'engageant activement dans la façon dont le monde est composé de codes ou de systèmes qui peuvent être piratés, de la "programmation, du langage, du langage poétique, des mathématiques ou de la musique, des courbes ou des colorations". De cette façon, le piratage peut être considéré comme une forme profonde d'alphabétisation axée sur la compréhension des différents systèmes et codes associés à la société.

Avec des hacks de haut niveau comme WikiLeaks et les fichiers Snowden suscitant une attention continue des médias, le piratage est de plus en plus considéré comme un moyen d'autonomiser les individus et de corriger les déséquilibres de pouvoir perçus. En tant que tels, les hackathons et les événements de programmation informatique en collaboration deviennent maintenant des activités régulières pour les groupes politiques et les militants aux côtés des communautés de programmeurs. L'industrie informatique et d'autres entreprises utilisent

également de plus en plus les approches de piratage. Le piratage n'est plus seulement pour les pirates, mais devient un moyen pour les entreprises technologiques de recruter de nouveaux employés ou d'externaliser simplement leur recherche et développement. Cependant, en tant que forme approfondie d'alphabétisation des données, le piratage et les hackathons attirent généralement ceux qui ont déjà des compétences techniques, tels que les passionnés d'informatique, les programmeurs et les concepteurs de logiciels[24].

- D. **Éducation aux médias et données personnelles** : ils ont tendance à suivre deux grandes approches, la première découle d'une approche de conception et se concentre sur la compétence des données et l'utilisation des données pour engager et autonomiser les individus dans la vie civique. Le second vise à aider les individus à comprendre et à manipuler les représentations de données sur les plateformes de médias sociaux.

La première approche suit la définition de la maîtrise des données au-delà de la littératie des données comme "le désir et la capacité de s'engager de manière constructive dans la société à travers les données". Cela a identifié quatre aspects sous-jacents de la littératie des données, notamment : l'éducation aux données, la visualisation des données, la modélisation des données et la participation aux données. Cette approche positionne la littératie des données comme un moyen d'identifier et de résoudre les problèmes du monde réel, qui sont de plus en plus médiés par les données. Étant donné que l'objectif de la maîtrise des données est d'aider les individus à apprendre à éclairer les phénomènes du monde réel grâce aux données, l'apprentissage doit être basé sur des projets, axé sur les problèmes et culturellement pertinent. Une compétence clé dans ce domaine de la littératie des données est ce que l'Oceans of Data Institute appelle la "pensée analytique", c'est-à-dire "prendre un problème (parfois) difficile, le décomposer et le reconstruire pour obtenir des informations intéressantes".

Le deuxième volet se concentre plus spécifiquement sur le rôle des données sur les plateformes de médias sociaux. Dans cette approche, la littératie des données implique de guider les individus à identifier, comprendre et manipuler les représentations des données en fonction de leurs besoins. Cette approche vise à aider les utilisateurs à décortiquer de manière critique la logique des médias sociaux et les normes, stratégies, économies et dynamiques qui sous-tendent les pratiques des médias sociaux. Ce volet de la littératie des données repose sur l'idée que les données sont "multivalentes", c'est-à-dire qu'elles ont de la valeur dans des régimes de valeurs multiples, parfois conflictuels. Par exemple, les données ont une valeur personnelle pour l'utilisateur, pour les opérateurs de plateforme qui ont pré-structuré des actes

de communication particuliers en données, et pour d'autres parties prenantes qui traitent et réutilisent ces données selon de nouveaux régimes de valeur. Les interventions d'alphabétisation des données se concentrent sur des cibles spécifiques telles que "identité", "activité", "interactivité" et "visibilité"[24].

### 1.3.3 Processus de la littératie des données

Le processus de la littératie des données divisée en cinq classes[24].

**Identification des données**, identification des données personnelles et de leur type.

**Compréhension des données**, identifier comment et où les données personnelles sont générées et traitées (traces de données), interpréter les informations représentées par les données traitées (visualisations de données, tableaux et graphiques).

**Réflexivité des données**, analyser et évaluer le profilage et les prédictions qui sont faites à partir des données personnelles traitées (c'est-à-dire l'analyse des sentiments, le traitement du langage naturel), comprendre les implications de la gestion, du contrôle et de l'application des données personnelles (critique individuelle et collective).

**Utilisations des données**, appliquer, gérer, contrôler les données, développer les compétences techniques et les compétences d'interprétation (lire les termes et conditions, ajuster les paramètres de confidentialité, bloquer les technologies, développer un langage partagé), appliquer les informations représentées par les données traitées (perspectives personnelles sur l'autonomie numérique et les performances )

**Tactiques des données**, utiliser des tactiques de résistance et d'obscurcissement (tactiques), réaffecter des données pour des raisons personnelles et sociales (applications créatives).

## 1.4 Classification des données personnelles

La classification des données est le processus d'organisation des données, "contenu, contexte, utilisateur" en fonction de leur utilisation afin qu'elles puissent être utilisées et protégées plus efficacement, nous pouvons les classer en plusieurs catégories comme suit :

**Identité** : c'est le fait d'être qui, le nom, l'âge, l'adresse, l'email, etc.

**Communication** : en communiquant avec un correspondant, est un courriel, enregistrement d'appel téléphonique de messagerie instantanée, échange de SMS, etc.

**Financier** : relevés bancaires (comptes personnels et conjoints), relevés de carte de crédit, contrats de logement / détails hypothécaires, etc.



- Familiale** : photographie, voyages, consommation d'énergie, calendriers partagés, les dossiers de santé, nombres des enfants, etc.
- Individuel** : traces de localisation personnelles, calendriers personnels, carnets d'adresses, données de suivi du sommeil, etc.
- Réseaux sociaux** : ce sont les candidats habituels (Twitter, Facebook, Instagram..) ainsi que ceux qui n'existent plus (msn).
- Voyages** : l'emplacement de déplacement, les lieux visités, réservation d'hôtel, etc.
- Démographie** : sur la population tels que la religion, la culture, la langue, le sexe, etc.
- Soins de santé** : les maladies, les traitements utilisés, le diagnostic des patients, etc.
- Événements de la vie** : sont des expériences comme le mariage, le divorce, la maladie ou une blessure, et le changement ou la perte d'un emploi, la mort, etc.

## 1.5 Protection des données personnelles

La protection des données est généralement définie comme la loi conçue pour protéger vos données personnelles. Dans les sociétés modernes, afin de nous permettre de contrôler nos données et de nous protéger contre les abus, il est essentiel que les lois sur la protection des données restreignent et façonnent les activités des entreprises et des gouvernements. Ces institutions ont montré à maintes reprises qu'à moins que des règles limitant leurs actions soient en place, elles s'efforceront de tout collecter, de tout miner, de tout garder, de le partager avec les autres, sans rien dire du tout[8].

### 1.5.1 Raisons morales pour protéger les données personnelles

Les raisons morales jouent un rôle important dans la protection des données, notamment les suivants :

- A. **Prévention des préjudices** : un accès illimité par des tiers à son compte bancaire, à son profil, à son compte de médias sociaux, à ses référentiels cloud, à ses caractéristiques et à sa localisation peut être utilisé pour nuire à la personne concernée de diverses manières.
- B. **Inégalité informationnelle** : les données personnelles sont devenues des marchandises. Les individus ne sont généralement pas en bonne position pour négocier des contrats sur l'utilisation de leurs données et n'ont pas les moyens de vérifier si les partenaires respectent les termes du contrat. Les lois, réglementations et gouvernance en matière de protection des données visent à établir des conditions équitables pour la rédaction de contrats de transmission et d'échange de données à

caractère personnel et à fournir aux personnes concernées des freins et contrepoids, des garanties de recours et des moyens de contrôler le respect des termes du contrat.

- C. **Injustice informationnelle et discrimination** : les informations personnelles fournies dans une sphère ou un contexte (par exemple, les soins de santé) peuvent changer de sens lorsqu'elles sont utilisées dans une autre sphère ou un autre contexte (comme les transactions commerciales) et peuvent entraîner une discrimination et des désavantages pour l'individu.
- D. **Empiètement sur l'autonomie morale et la dignité humaine** : le manque d'intimité peut exposer les individus à des forces extérieures qui influencent leurs choix et les amènent à prendre des décisions qu'ils n'auraient pas prises autrement. La surveillance de masse conduit à une situation où les individus font des choix et des décisions de manière routinière, systématique et continue parce qu'ils savent que d'autres les observent. Cela affecte leur statut d'êtres autonomes et à ce qui est parfois décrit comme un «effet dissuasif» sur eux et sur la société. Les considérations de violation du respect des personnes et de la dignité humaine sont étroitement liées[18].

### 1.5.2 Loi, réglementation et contrôle indirect sur l'accès

Reconnaissant qu'il existe des raisons morales de protéger les données personnelles, des lois sur la protection des données sont en vigueur dans presque tous les pays. Le principe moral de base qui sous-tend ces lois est l'exigence d'un consentement éclairé pour le traitement par la personne concernée, ce qui lui permet (au moins en principe) de contrôler les effets négatifs potentiels. En outre, le traitement des informations personnelles nécessite que leur objet soit spécifié, que leur utilisation soit limitée, que les individus soient informés et autorisés à corriger les inexactitudes, et que le détenteur des données soit responsable devant les autorités de contrôle. Parce qu'il est impossible de garantir la conformité de tous les types de traitement de données dans tous ces domaines et applications avec ces règles et lois de manière traditionnelle, les technologies dites *privacy-enhancing technologies*, «améliorant la confidentialité», (PET) et les systèmes de gestion d'identité devraient remplacer la surveillance humaine dans de nombreux cas. Le défi en matière de confidentialité au XXI<sup>e</sup> siècle est de garantir que la technologie est conçue de manière à intégrer les exigences de confidentialité dans le logiciel, l'architecture, l'infrastructure et les processus de travail de manière à ce que les violations de la vie privée ne se produisent pas. Les nouvelles générations de réglementations en matière de confidentialité (par exemple le RGPD Règlement général sur la protection des données) nécessitent désormais une approche standard de la «confidentialité par conception». Les écosystèmes de données et les systèmes socio-

techniques, les chaînes d'approvisionnement, les organisations, y compris les structures d'incitation, les processus commerciaux, le matériel et les logiciels techniques, la formation du personnel, devraient tous être conçus de telle manière que la probabilité de violations de la vie privée soit aussi faible que possible[18].

### 1.5.3 Données et contextes sensibles

Des normes plus strictes de protection des données devraient être appliquées dans le cadre de l'obtention, l'accès, la collecte, l'analyse ou autre utilisation de données concernant les populations vulnérables et les personnes à risque, les enfants et les jeunes, ou toutes autres données sensibles.

Il est important d'envisager la possibilité que le contexte puisse transformer des données non sensibles en données sensibles. Le contexte dans lequel les données sont utilisées (par exemple, les circonstances culturelles, géographiques, religieuses, politiques, etc.) peut avoir une incidence sur l'effet de l'analyse de ces données sur un ou plusieurs individus ou groupes d'individus, même si ces données ne revêtent pas un caractère explicitement personnel ou sensible[7].

### 1.5.4 La vulnérabilité

Le terme vulnérabilité définit une faiblesse sous-jacente associée à un système qui, si elle n'est pas corrigée à temps, expose le système à une menace potentielle[9].

- A. **Étapes de la vulnérabilité des données** : les données sont continuellement exposées aux menaces de cybersécurité en raison de plusieurs types de vulnérabilités qui se manifestent aux étapes suivantes :
- Au niveau du réseau** : clés USB, ordinateurs portables et netbooks (contiennent un port Ethernet pour accéder directement au réseau), des points d'accès sans fil, la personne non surveillée dans la salle des serveurs (les humains troyens), les ressources numériques, etc.
  - Au niveau du système** : phishing par e-mail, un utilisateur surveillant l'utilisateur pendant la connexion, les 3 programmes les plus exploités installés sur un système comprennent - Adobe Reader, Oracle Java, Adobe Flash Au niveau des données : appareils mobiles non sécurisés, systèmes de stockage cloud non sécurisés, etc.
  - Au niveau des données** : employés insatisfaits, appareils mobiles non sécurisés, applications de stockage, y compris le cloud, systèmes non sécurisés chez des fournisseurs de services tiers, pirates, etc[9].
- B. **Impact de la vulnérabilité des données** : la vulnérabilité des données à un impact sur le temps d'arrêt de l'entreprise, la perte de données et la confidentialité des données.

**Temps d'arrêt de l'entreprise :** le temps d'arrêt ou la panne se produit lorsqu'un système devient indisponible pendant une certaine durée et ne remplit pas sa fonction principale. Pour restaurer un système compromis à partir de zéro, l'entreprise doit investir des ressources, ce qui entraîne une perte initiale.

**Perte de données :** le chiffrement des données par un ransomware peut entraîner une perte permanente de données, compromettant ainsi l'avantage stratégique et affectant la réputation de la marque et la santé globale de l'entreprise. La prévention de la perte de données aurait pu être possible si les organisations avaient appliqué des correctifs en temps opportun.

**Confidentialité des données et implications juridiques :** l'accès non autorisé aux données tierces affecte la confidentialité, l'intégrité et la disponibilité des données organisationnelles, compromettant ainsi la confidentialité des données. Dans le contexte actuel, le non-respect des règles de confidentialité des données telles que le RGPD peut entraîner des complications juridiques[9].

### 1.5.5 Violation des données

Une violation de données est un incident de sécurité dans lequel des données sensibles, protégées ou confidentielles sont copiées, transmises, visualisées, volées ou utilisées par une personne non autorisée à le faire. Au fil du temps, des violations de données se sont produites dans divers secteurs, notamment le gouvernement, les soins de santé, les services financiers, les assurances, les médias sociaux, etc[20].

A. **Cause de violation de données :** la cause de la violation de données est un événement principal qui provoque un incident de violation de données. les causes peuvent être classés en intentionnelles et non intentionnelles[20].

- i Violation intentionnelle des données : un incident de violation causé par un acte malveillant dont l'intention est de nuire à une organisation.
  - Les pirates.
  - Initiés malveillants.
  - Acteurs parrainés par l'état.
  - Terroristes.
- ii Violation non intentionnelle de données : un incident de violation causé par des actions accidentelles d'un individu ou d'un processus et sans intention malveillante.
  - Comportement utilisateur non sécurisé.

- Perte ou réutilisation d'appareils multimédias.
- Logiciel défectueux.
- Divulgation non autorisée.
- Logiciels non autorisés.

**B. Impact de la violation des données :** l'impact d'une violation de données est l'effet négatif qu'un incident de violation de données peut avoir sur une organisation. Ces impacts peuvent être classés en impacts de confidentialité, de disponibilité et d'intégrité[20].

i Violation des données de confidentialité : les données sur les effets négatifs accessibles en dehors d'une exigence métier.

- Identité volée.
- Amendes.
- Poursuites.
- Perte d'avantage concurrentiel.
- Perte d'emploi.

ii Violation des données de disponibilité : toute perte d'accès aux données ou aux ressources de données pour une durée quelconque

- Déni de service.
- Données volées.
- Coupure de courant.
- Défaillance du système.
- Données supprimées.

iii Violation des données d'intégrité : toute manipulation non autorisée ou accidentelle des données au repos, en transit ou en cours d'utilisation

- Données modifiées.
- Perte d'avantage concurrentiel.
- Données supprimées.

**C. Résolutions de violation de données :** pour identifier les résolutions de violation de données, on s'appuie sur le cadre de gestion de la sécurité centré sur les incidents qui intègre les paradigmes de prévention et de réponse largement utilisés dans la gestion de la sécurité de l'information. Le paradigme de prévention est conçu pour éviter que des incidents de sécurité ne se produisent, tandis que le paradigme de réponse est destiné à réagir aux incidents de sécurité qui se sont produits. Sur la base des approches proactives et réactives du cadre, on a identifié trois catégories de résolution pour gérer les risques de violation de données : prévention, endiguement et récupération[20].

- i Prévention : une intervention visant à réduire la probabilité d'une violation de données et son impact négatif.
  - Soutien à la direction exécutive.
  - Gestion des politiques et des programmes.
  - Gestion des données.
  - Réseau sécurisé.
  - Gestion des identités et des accès.
  - Sensibilisation et formation.
  - Suivi.
  - Analyse comparative.
  - Évaluation des risques.
  - Test de pénétration.
- ii Endiguement : une intervention destinée à limiter l'ampleur et la portée d'une violation de données dès sa détection.
  - Système de détection et de prévention des incidents.
  - Équipe de sécurité informatique et de réponse aux incidents.
  - Suivi de l'attaquant.
  - Séparation du réseau.
- iii Récupération : une intervention visant à réduire l'impact négatif consécutif d'une violation de données après qu'elle s'est produite.
  - Assurance de cybersécurité.
  - Équipe de sécurité informatique et de réponse aux incidents.
  - Analyse des causes profondes.
  - Leçons apprises.
  - Résolution avant de reprendre les opérations.

### 1.5.6 Raison d'attaque

Le problème de la protection de la vie privée n'est pas nouveau et plusieurs attaques bien connues ont été identifiées dans la littérature. Ces attaques peuvent être classées comme suit :

**Attaque d'identification de l'utilisateur :** dont le but est d'exposer les requêtes d'identification de l'utilisateur sur une région spatiale, puis les requêtes ultérieures, plus spécifiques, impliquant des sous-régions spatiales de la requête originale.

**Suivi de localisation sensible :** où l'attaquant tente d'identifier un ou plusieurs emplacements que l'utilisateur visite fréquemment. Cette attaque peut révéler des informations sur la santé, le mode de vie, les habitudes d'une personne, etc.

**Attaque de suivi séquentiel :** qui vise à retrouver l'utilisateur en effectuant un ensemble de requêtes impliquant des régions spatio-temporelles adjacentes les unes aux autres puis en analysant les trajectoires de l'utilisateur pour identifier les lieux qu'il a visités[11].

### 1.5.7 Vie privée

La vie privée est au cœur des conceptions les plus élémentaires de la dignité humaine, et l'absence d'une définition convenue n'empêche pas le développement d'une compréhension large de la vie privée et de son importance dans une société démocratique. Les conceptions actuelles du droit à la vie privée réunissent trois aspects connexes : la vie privée décisionnelle, la vie privée informationnelle et la vie privée physique.

- Vie privée décisionnelle : une vision globale de la vie privée examine la capacité des individus à faire des choix de vie autonomes sans ingérence ni intimidation extérieures, y compris les conditions sociales, politiques et technologiques qui rendent cette "vie privée décisionnelle" possible. Cela fait de la vie privée une valeur sociale ainsi qu'un bien public et offre une protection contre les intrusions extérieures dans les foyers, les communications, les opinions, les croyances et les identités des gens.
- Vie privée informationnelle : la vie privée a évolué plus récemment pour englober un droit à la "vie privée informationnelle", également connu sous le nom de protection des données. Le droit à la confidentialité des informations est de plus en plus au cœur des politiques et des processus juridiques modernes et, dans la pratique, signifie que les individus devraient être en mesure de contrôler qui détient les données les concernant et quelles décisions sont prises sur la base de ces données.
- Vie privée physique : une troisième conception, plus simple, de la vie privée est celle de la "vie privée physique", le droit d'un individu à un espace privé et à l'intégrité corporelle[17].
- **Respecte de la vie privée :** un rapport du Rapporteur spécial au Conseil des droits de l'homme (A/HRC/23/40) définit le respect de la vie privée comme "la présomption selon laquelle les individus devraient disposer d'un domaine de développement autonome, d'échanges et de liberté, une "sphère privée" avec ou sans interaction avec autrui, libre de toute intervention de l'État et ingérence excessive non sollicitée d'autres individus" [7].
- **L'impact des technologies de l'information sur la vie privée,** parmi les impacts de la technologie sur la vie privée :
  - i Évolution des technologies de l'information : la «technologie de l'information» fait référence aux systèmes automatisés de stockage, de traitement et de distribution d'informations. Cela implique généralement l'utilisation d'ordinateurs et de réseaux de

communication. La quantité d'informations pouvant être stockées ou traitées dans un système d'information dépend de la technologie utilisée. La capacité de la technologie a augmenté rapidement au cours des dernières décennies. Cela vaut pour la capacité de stockage, la capacité de traitement et la bande passante de communication. Nous sommes désormais capables de stocker et de traiter des données au niveau de l'exaoctet.

- ii Internet : l'internet, conçu à l'origine dans les années 60 et développé dans les années 80 comme un réseau scientifique d'échange d'informations, n'a pas été conçu dans le but de séparer les flux d'informations. Le World Wide Web d'aujourd'hui n'était pas prévu, pas plus que la possibilité d'une utilisation abusive d'Internet.
- iii Les médias sociaux : posent des défis supplémentaires. La question ne porte pas seulement sur les raisons morales de limiter l'accès à l'information, mais aussi sur les raisons morales de limiter les invitations aux utilisateurs à soumettre toutes sortes d'informations personnelles. Les sites de réseaux sociaux invitent l'utilisateur à générer plus de données, pour augmenter la valeur du site.
- iv BigData : les utilisateurs génèrent des charges de données lorsqu'ils sont en ligne. Il s'agit non seulement de données saisies explicitement par l'utilisateur, mais également de nombreuses statistiques sur le comportement des utilisateurs : sites visités, liens cliqués, termes de recherche saisis, etc. L'exploration de données peut être utilisée pour extraire des modèles de ces données, qui peuvent ensuite être utilisées pour créer des décisions concernant l'utilisateur. Celles-ci peuvent uniquement affecter l'expérience en ligne (publicités affichées), mais, selon les parties qui ont accès aux informations, elles peuvent également avoir un impact sur l'utilisateur dans des contextes complètement différents.
- v Appareils mobiles : comme les utilisateurs possèdent de plus en plus d'appareils en réseau tels que les téléphones intelligents, les appareils mobiles collectent et envoient de plus en plus de données. Ces appareils contiennent généralement une gamme de capteurs générant des données, y compris le GPS (localisation), des capteurs de mouvement et des caméras, et peuvent transmettre les données résultantes via Internet ou d'autres réseaux.
- vi L'internet des objets : les appareils connectés à Internet ne se limitent pas aux appareils informatiques appartenant à l'utilisateur comme les smartphones. De nombreux appareils contiennent des puces et / ou sont connectés au soi-disant Internet des objets. Les puces RFID (identification par radiofréquence) peuvent être lues à une distance limitée, de sorte que vous pouvez les tenir devant



un lecteur plutôt que de les insérer. Les passeports de l'UE et des États-Unis ont des puces RFID avec des données biométriques protégées, mais des informations telles que la nationalité de l'utilisateur peuvent facilement fuir lorsque vous essayez de lire de tels appareils). Les RFID «intelligents» sont également intégrés dans les systèmes de paiement des transports publics. Les RFID «stupides», ne contenant essentiellement qu'un numéro, apparaissent dans de nombreux types de produits en remplacement du code-barres et pour une utilisation en logistique. Pourtant, ces puces pourraient être utilisées pour retracer une personne une fois que l'on sait qu'elle porte un article contenant une puce.

- vii Gouvernement : le gouvernement et l'administration publique ont également subi des transformations radicales du fait de la disponibilité de systèmes informatiques avancés. Des exemples de ces changements sont les passeports biométriques, les services en ligne de gouvernement électronique, les systèmes de vote, une variété d'outils et de plateformes de participation des citoyens en ligne ou l'accès en ligne aux enregistrements des sessions du Parlement et des réunions des commissions gouvernementales[18].

### 1.5.8 Confidentialité des données

La confidentialité des données ou la confidentialité des informations est une branche de la sécurité des données qui concerne le traitement approprié des données - consentement, notification et obligations réglementaires. Plus précisément, les problèmes pratiques de confidentialité des données tournent souvent autour de :

- Si ou comment les données sont partagées avec des tiers.
- Comment les données sont collectées ou stockées légalement.
- Restrictions réglementaires telles que GDPR, HIPAA ou GLBA.

**L'importance de la confidentialité des données :** il existe deux facteurs expliquant pourquoi la confidentialité des données est l'un des problèmes les plus importants.

- A. Les données sont l'un des actifs les plus importants d'une entreprise. Avec l'essor de l'économie des données, les entreprises trouvent une énorme valeur dans la collecte, le partage et l'utilisation des données. Des entreprises telles que Google, Facebook et Amazon ont toutes bâti des empires au sommet de l'économie des données. La transparence dans la façon dont les entreprises demandent le consentement, se conforment à leurs politiques de confidentialité et gèrent les données qu'elles ont collectées est essentielle pour instaurer la confiance et la

responsabilité avec les clients et partenaires qui attendent la confidentialité. De nombreuses entreprises ont appris l'importance de la vie privée à la dure, grâce à des échecs de confidentialité très médiatisés.

- B. La vie privée est le droit d'un individu d'être à l'abri d'une surveillance non sollicitée. Exister en toute sécurité dans son espace et exprimer librement ses opinions à huis clos est essentiel pour vivre dans une société démocratique[25].

## 1.6 Conclusion

Dans ce chapitre, nous avons vu un aperçu de la protection des données personnelles, dans le deuxième chapitre, nous présentons la sécurité, la technologie blockchain, certains travaux connexes et le synthèse.

# Chapitre 2

## L'état de l'art

---

*"SHA-256 est très solide. Ce n'est pas comme l'étape incrémentielle de MD5 à SHA1. Cela peut durer plusieurs décennies à moins qu'il y ait une attaque de percée massive."*

*Satoshi Nakamoto*

---

### 2.1 Introduction

La sécurité des données protège les données personnelles contre les intrusions et garantit l'intégrité et la confidentialité à travers l'utilisation des méthodes et techniques de sécurité.

Dans ce chapitre, nous présentons la sécurité, ses services, la cryptographie et ses techniques, la cryptanalyse, les fonctions de hachage, puis nous parlons de la blockchain, son historique, ses composants, les modèles de consensus, les contrats intelligents, et enfin la blockchain et la conservation des données ainsi que une explication pour certains travaux qui utilisent la blockchain pour la conservation des données.

### 2.2 Sécurité

En général, la sécurité est la qualité ou l'état de sécurité. En d'autres termes, la protection contre les adversaires contre ceux qui feraient du mal intentionnellement ou non. La sécurité nationale, par exemple, est un système à plusieurs niveaux qui protège la souveraineté d'un État, ses actifs, ses ressources et son peuple. Atteindre le niveau de sécurité approprié pour une organisation nécessite également un système multiforme.

Une organisation performante doit disposer des multiples niveaux de sécurité suivants pour protéger ses[6] :

- Sécurité physique, pour protéger les objets, objets ou zones physiques contre tout accès non autorisé et toute utilisation abusive.
- Sécurité du personnel, pour protéger la personne ou le groupe de personnes autorisées à accéder à l'organisation et à ses opérations.
- Sécurité des opérations, pour protéger les détails d'une opération particulière ou d'une série d'activités.
- Sécurité des communications, pour protéger les supports de communication, la technologie et le contenu.
- Sécurité réseau, pour protéger les composants réseau, les connexions et le contenu.
- Sécurité de l'information, pour protéger la confidentialité, l'intégrité et la disponibilité des actifs informationnels, qu'ils soient stockés, traités ou transmis. Il est atteint grâce à l'application des politiques, de l'éducation, de la formation et de la sensibilisation et de la technologie.

### 2.2.1 Service de sécurité

Un service qui améliore la sécurité des systèmes de traitement des données et les transferts d'informations d'une organisation. Les services sont destinés à contrer les attaques de sécurité et ils utilisent un ou plusieurs mécanismes de sécurité pour fournir le service. La classification des services de sécurité est la suivante[12] :

**Confidentialité** : garantit que les informations contenues dans un système informatique et les informations transmises ne sont accessibles qu'à la lecture par les parties autorisées. Par exemple. Impression, affichage et autres formes de divulgation.

**Authentication** : garantit que l'origine d'un message ou d'un document électronique est correctement identifiée, avec l'assurance que l'identité n'est pas fausse.

**Intégrité** : garantit que seules les parties autorisées peuvent modifier les actifs du système informatique et les informations transmises. La modification comprend l'écriture, le changement de statut, la suppression, la création et le retard ou la relecture des messages transmis.

**Non de répudiation** : nécessite que ni l'expéditeur ni le destinataire d'un message ne puissent refuser la transmission.

**Disponibilité** : nécessite que les actifs du système informatique soient disponibles pour les parties autorisées en cas de besoin.

### 2.2.2 Cryptographie

La cryptographie est l'art du codage secret qui est utilisée depuis l'époque romaine pour cacher des informations secrètes ou sécuriser un message. Le service de base fourni par la cryptographie est la possibilité d'envoyer les informations entre les participants d'une manière qui empêche les autres de les lire.

Le but principal de la cryptographie est utilisé non seulement pour assurer la confidentialité, mais aussi pour fournir des solutions à d'autres problèmes tels que : l'intégrité des données, l'authentification, la non répudiation, le contrôle d'accès.

Pour garder les informations secrètes, une méthode largement utilisée est le cryptage / décryptage, sont les fonctions fondamentales de la cryptographie. En cryptage, un simple message (texte brut) est converti en une forme illisible appelée texte chiffré. Pendant le décryptage, un texte chiffré est converti en texte d'origine (texte en clair). Ces deux fonctions sont utilisées pour sécuriser le message contre qui n'est pas autorisé à afficher le contenu du message. Le fonctionnement simple des fonctions de chiffrement et de déchiffrement est illustré à la figure 2.1 [26][23].

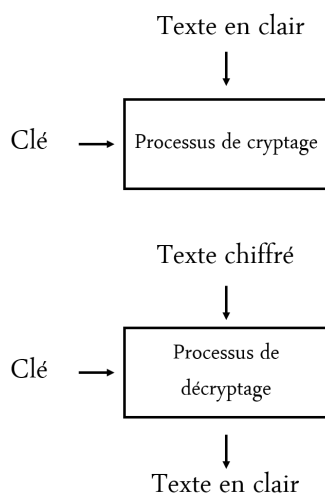


FIGURE 2.1: Processus de cryptographie[26]

### 2.2.3 Techniques de cryptographie

Il existe deux techniques de base pour crypter les informations : le cryptage symétrique également appelé cryptage à clé secrète et le cryptage asymétrique également appelé cryptage à clé publique.

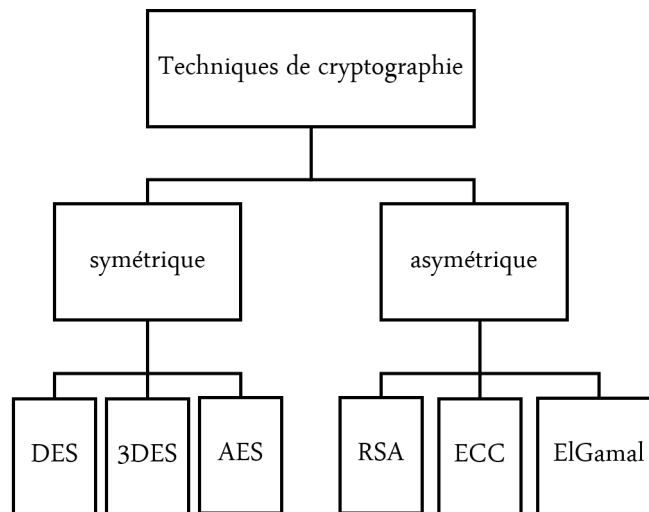


FIGURE 2.2: Techniques de cryptographie[23]

### Cryptographie symétrique

La cryptographie symétrique est placée dans la catégorie des schémas de cryptographie dans lesquels une clé partagée est utilisée pour convertir un texte en clair en texte chiffré. Une même clé secrète est partagée par l'expéditeur et le destinataire[23].

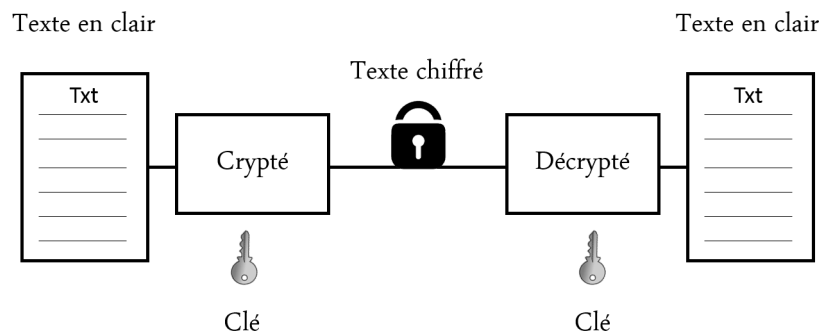


FIGURE 2.3: La cryptographie symétrique

**DES** (Data Encryption Standard) : DES signifie Data Encryption Standard. DES introduit au début des années 1970 chez IBM. La première conception de DES est basée sur Horst Feistel. DES est un algorithme cryptographique symétrique utilisé pour le chiffrement et le déchiffrement des messages. Dans DES, une seule clé secrète est utilisée à la fois pour le chiffrement et le déchiffrement. La taille de clé de DES est de 56 bits. Pour effectuer le chiffrement / déchiffrement,

l'expéditeur et le destinataire doivent avoir la même clé. Le DES effectue le chiffrement sur un bloc de 64 bits. L'algorithme DES est le plus largement utilisé dans de nombreuses applications et certaines utilisations populaires dans les systèmes militaires, commerciaux et de sécurité des communications.

**3DES** identique à DES mais la taille de la clé est différente de DES. La taille de clé de 3DES est de 168 bits. L'algorithme 3DES effectue l'opération trois fois sur chaque bloc de données. Il est plus lent que DES.

**AES** (Advanced Encryption Standard) : AES signifie Advanced Encryption Standard qui est l'avancement de l'algorithme 3DES. Il a été introduit en 1997 par le NIST (National Institute of Standards and Technology). Fondamentalement, AES est basé sur le chiffrement Rijndael développé par deux cryptographes, Joan Daemen et Vincent Rijmen. AES est différent de DES et 3DES en raison des tailles de clés variables telles que 128, 192 et 256 bits. Comme DES et 3DES, AES effectue également le chiffrement sur des blocs de 128 bits. Algorithme AES utilisé dans les petits appareils pour crypter un message à envoyer sur un réseau. Certaines autres applications sont des transactions monétaires et des applications de sécurité.

### Cryptographie asymétrique

La cryptographie asymétrique fait également partie de la catégorie des schémas de cryptographie. Contrairement à la cryptographie symétrique, deux clés sont utilisées : l'une est publique et la seconde est privée. La clé publique est partagée par quiconque dans le système cryptographique tandis que la clé privée est gardée secrète par l'utilisateur authentifié[23].

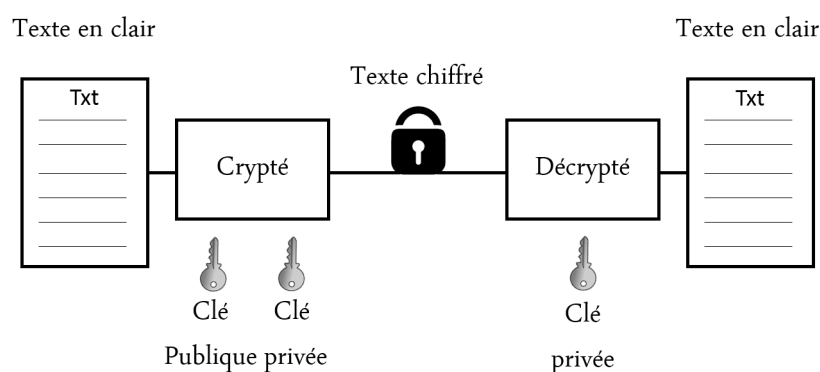


FIGURE 2.4: La cryptographie asymétrique

**RSA** (Rivest, Shamir et Adleman) : RSA signifie Rivest, Shamir et Adleman qui a introduit l'algorithme RSA en 1977. RSA est un algorithme

cryptographique asymétrique qui est également utilisé pour le cryptage et le décryptage du message. RSA est largement utilisé dans le transfert de clés sur un canal non sécurisé. En raison de la nature asymétrique, l'algorithme utilise deux clés. L'une est la clé publique et la seconde est une clé privée. La clé publique est accessible à tous dans le cryptosystème et la clé privée est gardée secrète par une personne autorisée. RSA assure la confidentialité, l'intégrité, l'authenticité et la non-répudiation des données. RSA est plus couramment utilisé dans l'industrie électronique pour le transfert d'argent en ligne.

**ECC** (Elliptic Curve Cryptography) : ECC signifie Elliptic Curve Cryptography. ECC introduit en 1985 par Neal Koblitz et Victor S. Miller. ECC appartient à la catégorie du schéma asymétrique basé sur des courbes elliptiques. Les applications de l'ECC sont le cryptage, les signatures numériques et les générateurs pseudo-aléatoires.

**ElGamal** L'algorithme ElGamal a été introduit en 1985 par Taher ElGamal. ElGamal est un algorithme de chiffrement de clé asymétrique basé sur l'échange de clés Diffie-Helman comme alternative à RSA pour le chiffrement à clé publique. ElGamal est également utilisé dans l'algorithme de génération de signature numérique appelé schéma de signature ElGamal.

#### 2.2.4 Cryptanalyse

l'étude des principes et des méthodes de transformation d'un message inintelligible en un message intelligible sans connaissance de la clé. Également appelé rupture de code[12].

#### 2.2.5 Fonction de hachage cryptographique

Une fonction de hachage cryptographique  $H$  est une fonction qui prend en entrée des chaînes de bits de longueur arbitraire et produit en sortie une chaîne de bits de longueur fixe ; la sortie est souvent appelée résumé, hashcode ou valeur de hachage. Les fonctions de hachage sont beaucoup utilisées en informatique, mais la différence cruciale entre une fonction de hachage standard et une fonction de hachage cryptographique est qu'une fonction de hachage cryptographique devrait au moins avoir la propriété d'être à sens unique. En d'autres termes, étant donné toute chaîne  $y$  du domaine de codage de  $H$ , il devrait être impossible de trouver une valeur  $x$  dans le domaine de  $H$  telle que  $H(x) = y$ [27].

Propriétés des fonctions de hachage[28] :

- résistant à la pré-image : cela signifie qu'ils sont à sens unique, il est impossible de calculer la valeur d'entrée correcte compte tenu d'une certaine valeur de sortie (par exemple, étant donné un digest, trouvez  $x$ , tel que  $\text{hachage}(x) = \text{digest}$ ).



- seconde préimage résistante : cela signifie que l'on ne peut pas trouver une entrée hachée sur une sortie spécifique. Plus précisément, les fonctions de hachage cryptographiques sont conçues de sorte que, étant donné une entrée spécifique, il est impossible de trouver une deuxième entrée qui produise la même sortie (par exemple, étant donné  $x$ , trouver  $y$  tel que  $\text{hachage}(x) = \text{hachage}(y)$ ).

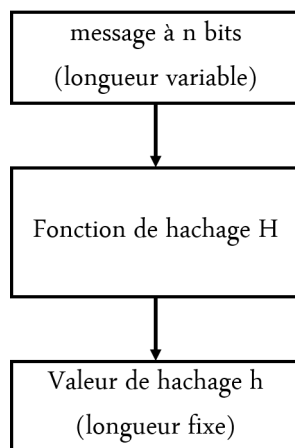


FIGURE 2.5: Schéma fonctionnel de la fonction de hachage[21]

Les fonctions de hachage les plus largement déployées sont MD-5, RIPEMD-160, SHA-1 et SHA-2, qui sont toutes basées sur la construction Merkle – Damg usinard utilisant une fonction de compression fixe (c'est-à-dire sans clé)  $f$ . L'algorithme MD-5 produit des sorties de 128 bits, tandis que RIPEMD 160 et SHA-1 produisent tous deux des sorties de 160 bits, tandis que SHA-2 est en fait trois algorithmes, SHA-256, SHA-384 et SHA-512, ayant des sorties de 256, 384 et 512 bits respectivement. Toutes ces fonctions de hachage sont dérivées d'un algorithme plus simple antérieur appelé MD-4. Les sept algorithmes principaux de la famille MD-4 sont[27] :

- **MD-4** : La fonction  $f$  a 3 tours de 16 étapes et une longueur de bit de sortie de 128 bits.
- **MD-5** : La fonction  $f$  a 4 tours de 16 étapes et une longueur de bit de sortie de 128 bits.
- **SHA-1** : La fonction  $f$  a 4 tours de 20 étapes et une longueur de bit de sortie de 160 bits.
- **RIPEMD-160** : La fonction  $f$  a 5 tours de 16 étapes et une longueur de bit de sortie de 160 bits.
- **SHA-256** : la fonction  $f$  a 64 tours de pas simples et une longueur de bit de sortie de 256 bits.

- **SHA-384** : La fonction  $f$  est identique à SHA-512 sauf que la sortie est tronquée à 384 bits et que la valeur de chaînage initiale  $H$  est différente.
  - **SHA-512** : la fonction  $f$  a 80 tours de pas simples et une longueur de bit de sortie de 512 bits.
- SHA** : Secure Hash Algorithm est la fonction de hachage la plus connue et la plus utilisée, ainsi qu'il y'a des différents paramètres SHA est présentée dans le tableau 2.1. Toutes les tailles sont mesurées en bits.

| Algorithme | Taille de digest | Taille de message | Taille de block | Taille de mot | pas de pas |
|------------|------------------|-------------------|-----------------|---------------|------------|
| SHA-1      | 160              | $<2^{64}$         | 512             | 32            | 80         |
| SHA-224    | 224              | $<2^{64}$         | 512             | 32            | 64         |
| SHA-256    | 256              | $<2^{64}$         | 512             | 32            | 64         |
| SHA-384    | 384              | $<2^{128}$        | 1024            | 64            | 80         |
| SHA-512    | 512              | $<2^{128}$        | 1024            | 64            | 80         |

TABLE 2.1: Comparaison des paramètres SHA[21]

## 2.3 Blockchain

Est une technologie de stockage et de transmission d'information qui est sécurisé et transparente et qui fonctionne sans organe central de contrôle.

### 2.3.1 Historique

Les idées fondamentales derrière la technologie blockchain sont apparues à la fin des années 1980 et au début des années 1990. En 1989, Leslie Lamport a développé le protocole Paxos, et en 1990 a soumis le document Le Parlement à temps partiel à ACM Transactions on Computer Systems, le document a finalement été publié dans un numéro de 1998. L'article décrit un modèle de consensus pour parvenir à un accord sur un résultat dans un réseau d'ordinateurs où les ordinateurs ou le réseau lui-même peuvent ne pas être fiables. En 1991, une chaîne d'informations signée a été utilisée comme registre électronique pour signer numériquement des documents d'une manière qui pouvait facilement montrer qu'aucun des documents signés de la collection n'avait été modifié. Ces concepts ont été combinés et appliqués à la monnaie électronique en 2008 et décrits dans le document Bitcoin : A Peer to Peer Electronic Cash System, publié sous un pseudonyme par Satoshi Nakamoto, puis plus tard en 2009 avec la création du réseau de blockchain de crypto-monnaie Bitcoin. Le document de Nakamoto contenait le plan que la plupart des schémas de crypto-monnaie modernes suivent (bien qu'avec des variations et des modifications). Bitcoin n'était que la première des nombreuses applications de blockchain[28].

### 2.3.2 Composants de blockchain

La technologie de la blockchain utilise des mécanismes informatiques et des primitives cryptographiques bien connus (fonctions de hachage cryptographique, signatures numériques, cryptographie à clé asymétrique)[28], ainsi que les principaux composants de la blockchain sont : fonctions de hachage cryptographiques, transactions, cryptographie à clé asymétrique, adresses, registres, blocs, validateurs, chaînage des blocs.

- A. **Le hachage** : est une méthode d'application d'une fonction de hachage cryptographique aux données, qui calcule une sortie relativement unique (appelée digest message, ou simplement digest) pour une entrée (un fichier, texte ou une image)[28]. Le digest est toujours le même pour certaine entrée mais la modification de l'entrée donne un résultat complètement différent comme l'exemple de tableau 2.2.

| Input Text    | SHA-256 Digest Value   |
|---------------|--|
| 1             | 0x6b86b273ff34fce19d6b804eff5a3f5747ada4eaa22f1d49c01e52ddb7875b4b |
| 2             | 0xd4735e3a265e16eee03f59718b9b5d03019c07d8b6c51f90da3a666eec13ab35 |
| Hello, World! | 0xdffd6021bb2bd5b0af676290809ec3a53191dd81c7f70a4b28688a362182986f |

TABLE 2.2: Valeurs de digest SHA-256 correspondantes à le texte entré[28]

**Nonce cryptographique** : Un nonce cryptographique est un nombre arbitraire qui n'est utilisé qu'une seule fois. Un nonce cryptographique peut être combiné avec des données pour produire différents digests de hachage par nonce :

$$\text{hash}(\text{data} + \text{nonce}) = \text{digest}$$

Seule la modification de la valeur nonce fournit un mécanisme permettant d'obtenir différentes valeurs de digest tout en conservant les mêmes données. Cette technique est utilisée dans le modèle de consensus de preuve de travail[28].

- B. **Les transactions** : une transaction représente une interaction entre les parties. Avec les crypto-monnaies comme la figure 2.6 montre, une transaction représente un transfert de la crypto-monnaie entre les utilisateurs du réseau blockchain. Pour les scénarios d'entreprise à entreprise, une transaction peut être un moyen d'enregistrer des activités se produisant sur des actifs numériques ou physiques, la figure 2.7 présente un exemple de transactions entre un médecin et un patient, chaque bloc d'une blockchain peut contenir zéro ou plusieurs transactions[28].

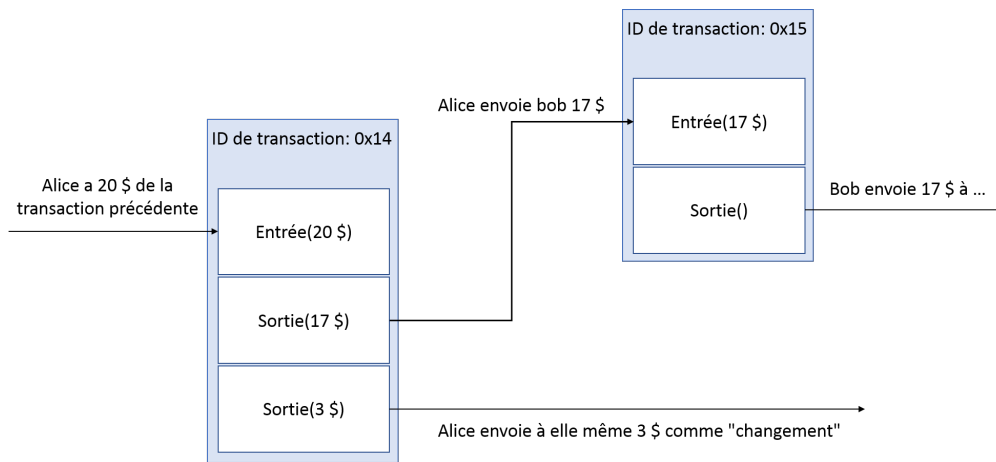


FIGURE 2.6: Exemple de transaction de crypto-monnaie[28]

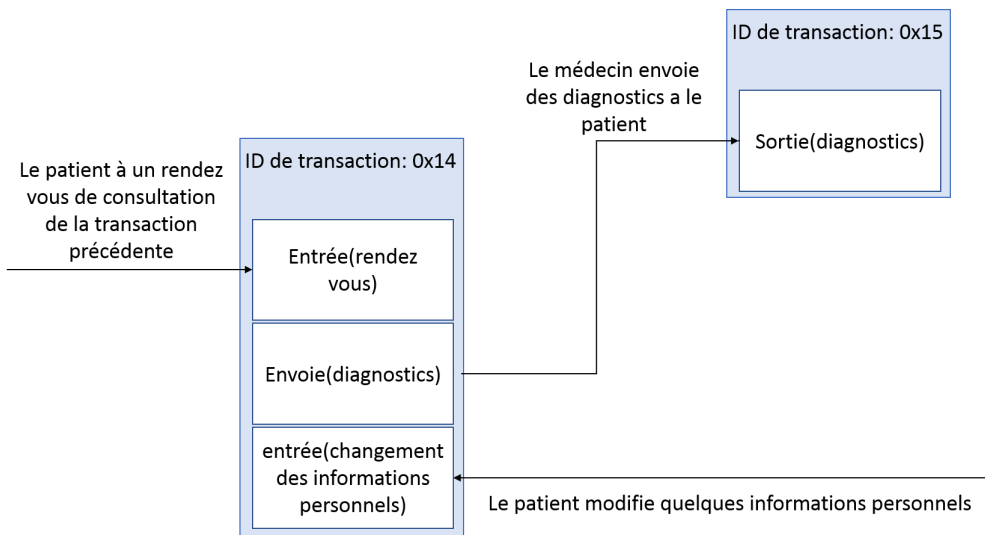


FIGURE 2.7: Exemple de transaction des activités entre le médecin et le patient

**C. La cryptographie asymétrique :** l'utilisation de la cryptographie asymétrique en blockchain est mesurée aux points suivants[28] :

- Les clés privées sont utilisées pour signer numériquement les transactions.
- Les clés publiques sont utilisées pour dériver des adresses.
- Les clés publiques sont utilisées pour vérifier les signatures générées avec des clés privées.

- La cryptographie à clé asymétrique permet de vérifier que l'utilisateur qui transfère de la valeur à un autre utilisateur est en possession de la clé privée capable de signer la transaction.

D. **Les adresses et dérivation d'adresses** : certains réseaux de blockchain utilisent une adresse, qui est une courte chaîne alphanumérique de caractères dérivée de la clé publique de l'utilisateur du réseau de blockchain utilisant une fonction de hachage cryptographique, ainsi que des données supplémentaires (par exemple, numéro de version, sommes de contrôle). La plupart des implémentations de blockchain utilisent des adresses comme points de terminaison «vers» et «depuis» dans une transaction. Les adresses sont plus courtes que les clés publiques et ne sont pas secrètes. Une méthode pour générer une adresse consiste à créer une clé publique, en lui appliquant une fonction de hachage cryptographique et en convertissant le hachage en texte :

Clé publique  $\rightarrow$  fonction de hachage cryptographique  $\rightarrow$  adresse

Les adresses peuvent agir comme identifiant accessible au public dans un réseau de blockchain pour un utilisateur, et souvent une adresse sera convertie en un code QR (Quick Response Code, un code à barres bidimensionnel qui peut contenir des données arbitraires) pour une utilisation plus facile avec le mobile dispositifs[28].



FIGURE 2.8: Exemple de code QR[28]

**Stockage de clé privée** : Avec certains réseaux de chaînes de blocs, les utilisateurs doivent gérer et stocker en toute sécurité leurs propres clés privées. Au lieu de les enregistrer manuellement, ils utilisent souvent un logiciel pour les stocker en toute sécurité. Ce logiciel est souvent appelé portefeuille. Le portefeuille peut stocker des clés privées, des clés publiques et des adresses associées. Il peut également exécuter d'autres fonctions, telles que le calcul du nombre total de ressources numériques qu'un utilisateur peut posséder. Si un utilisateur perd une clé privée, tout actif numérique associé à cette clé est perdu, car il est impossible de régénérer la même clé privée sur le plan informatique. Si une clé privée est volée, l'attaquant aura un accès complet à tous les actifs numériques contrôlés par cette clé privée[28].

E. **les registres** : un registre est un ensemble de transactions. Tout au long de l'histoire, des registres à stylo et papier ont été utilisés pour suivre l'échange de biens et de services. Dans les temps modernes, les registres ont été stockés numériquement, souvent dans de grandes bases de données détenues et exploitées par un tiers de confiance centralisé (c'est-à-dire le propriétaire du registre) au nom d'une communauté d'utilisateurs. Ces registres avec propriété centralisée peuvent être mis en œuvre de manière centralisée ou distribuée (c'est-à-dire, un seul serveur ou un cluster de serveurs de coordination).

Il est de plus en plus intéressant d'explorer la possibilité de répartir la propriété du registre. La technologie Blockchain permet une telle approche en utilisant à la fois la propriété distribuée et une architecture physique distribuée. L'architecture physique distribuée des réseaux de chaînes de blocs implique souvent un ensemble d'ordinateurs beaucoup plus important que celui typique d'une architecture physique distribuée gérée de manière centralisée. L'intérêt croissant pour la propriété répartie des registres est dû à des problèmes de confiance, de sécurité et de fiabilité liés aux registres à propriété centralisée[28] :

- Les registres appartenant à un organisme central peuvent être perdus ou détruits, un utilisateur doit avoir confiance que le propriétaire sauvegarde correctement le système.
  - Un réseau blockchain est distribué par conception, créant de nombreuses copies de sauvegarde, toutes mises à jour et synchronisées avec les mêmes données de registre entre pairs. Un avantage clé de la technologie blockchain est que chaque utilisateur peut conserver sa propre copie du registre. Chaque fois que de nouveaux nœuds complets rejoignent le réseau de la blockchain, ils tentent de découvrir d'autres nœuds complets et demandent une copie complète du registre du réseau de la blockchain, ce qui rend la perte ou la destruction du blockchain difficile.
- Les registres détenus de manière centralisée peuvent se trouver sur un réseau homogène, où tous les logiciels, le matériel et l'infrastructure réseau peuvent être identiques. En raison de cette caractéristique, la résilience globale du système peut être réduite car une attaque sur une partie du réseau fonctionnera partout.
  - Un réseau blockchain est un réseau hétérogène, où les logiciels, le matériel et l'infrastructure réseau sont tous différents. En raison des nombreuses différences entre les nœuds sur le réseau de la chaîne de blocs, une attaque sur un nœud n'est pas garantie de fonctionner sur d'autres nœuds.

- Les registres appartenant à l'administration centrale peuvent être entièrement situés dans des emplacements géographiques spécifiques (par exemple, tous dans un seul pays). Si des pannes de réseau devaient se produire à cet endroit, le registre et les services qui en dépendent pourraient ne pas être disponibles.

Un réseau de blockchain peut être composé de nœuds géographiquement divers qui peuvent être trouvés dans le monde. De ce fait, et le réseau de chaînes de blocs fonctionnant de manière poste à poste, il résiste à la perte de n'importe quel nœud, voire d'une région entière de nœuds.

- Les transactions sur un registre central ne sont pas effectuées de manière transparente et peuvent ne pas être valides ; un utilisateur doit avoir confiance que le propriétaire valide chaque transaction reçue.
  - Un réseau de blockchain doit vérifier que toutes les transactions sont valides ; si un nœud malveillant transmettait des transactions non valides, d'autres les détecteraient et les ignoreraient, empêchant les transactions non valides de se propager à travers le réseau de la chaîne de blocs.
- La liste des transactions d'un registre appartenant à l'administration centrale peut ne pas être complète ; un utilisateur doit avoir confiance que le propriétaire inclut toutes les transactions valides qui ont été reçues.
  - Un réseau de blockchain détient toutes les transactions acceptées dans son registre distribué. Pour construire un nouveau bloc, une référence doit être faite à un bloc précédent - donc en construisant dessus. Si un nœud de publication n'incluait pas de référence au dernier bloc, les autres nœuds le rejetteraient.
- Les données de transaction d'un registre appartenant à l'administration centrale peuvent avoir été modifiées ; un utilisateur doit avoir confiance que le propriétaire ne modifie pas les transactions passées.
  - Un réseau de chaînes de blocs utilise des mécanismes cryptographiques tels que des signatures numériques et des fonctions de hachage cryptographiques pour fournir des registres inviolables et inviolables.
- Le système centralisé peut ne pas être sûr ; un utilisateur doit avoir confiance que les systèmes et réseaux informatiques associés reçoivent des correctifs de sécurité critiques et ont mis en

œuvre les meilleures pratiques de sécurité. Le système peut être piraté et des informations personnelles ont été volées en raison de l'insécurité.

- Un réseau blockchain, en raison de la nature distribuée, ne fournit aucun point d'attaque centralisé. Généralement, les informations sur un réseau de blockchain sont visibles publiquement et n'offrent rien à voler. Pour attaquer les utilisateurs du réseau blockchain, un attaquant devrait les cibler individuellement. Le ciblage de la blockchain elle-même se heurterait à la résistance des nœuds honnêtes présents dans le système. Si un nœud individuel n'était pas corrigé, cela n'affecterait que ce nœud - pas le système dans son ensemble.

F. **Les blocs** : les utilisateurs du réseau Blockchain soumettent des transactions candidates au réseau Blockchain via un logiciel (applications de bureau, applications pour smartphone, portefeuilles numériques, services Web, etc.). Le logiciel envoie ces transactions à un ou plusieurs nœuds du réseau blockchain. Les nœuds choisis peuvent être des nœuds complets non publieurs ainsi que des nœuds de publication. Les transactions soumises sont ensuite propagées aux autres nœuds du réseau, mais cela en soi ne place pas la transaction dans la blockchain. Pour de nombreuses implémentations de blockchain, une fois qu'une transaction en attente a été distribuée aux nœuds, elle doit ensuite attendre dans une file d'attente jusqu'à ce qu'elle soit ajoutée à la blockchain par un nœud de publication.

Les transactions sont ajoutées à la blockchain lorsqu'un nœud de publication publie un bloc. Un bloc contient un en-tête de bloc et des données de bloc. L'en-tête du bloc contient des métadonnées pour ce bloc. Les données de bloc contiennent une liste de transactions validées et authentiques qui ont été soumises au réseau de blockchain. La validité et l'authenticité sont garanties en vérifiant que la transaction est correctement formatée et que les fournisseurs d'actifs numériques dans chaque transaction (répertoriés dans les valeurs «d'entrée» de la transaction) ont chacun signé la transaction de manière cryptographique. Cela vérifie que les fournisseurs d'actifs numériques pour une transaction avaient accès à la clé privée qui pouvait signer les actifs numériques disponibles. Les autres nœuds complets vérifieront la validité et l'authenticité de toutes les transactions dans un bloc publié et n'accepteront pas un bloc s'il contient des transactions non valides.

Il convient de noter que chaque implémentation de la blockchain peut définir ses propres champs de données; cependant, de nombreuses implémentations de blockchain utilisent des champs de données comme les suivants[28] :

- En-tête de bloc



- Le numéro de bloc, également connu sous le nom de hauteur de bloc dans certains réseaux de chaînes de blocs.
  - La valeur de hachage de l'en-tête de bloc précédent.
  - Une représentation de hachage des données de bloc (différentes méthodes peuvent être utilisées pour accomplir cela, et stocker le hachage racine, ou en utilisant un hachage de toutes les données de bloc combinées).
  - Un horodatage.
- Bloc des données
    - Une liste des transactions et des événements du grand livre inclus dans le bloc.
    - D'autres données peuvent être présentes.

G. **Validateurs** : est une personne chargée de vérifier les transactions au sein d'une blockchain. Il est important de noter que «validation» et «consensus» ne sont pas la même chose. Un validateur Blockchain effectue la validation en vérifiant que les transactions sont légales (non malveillantes, doubles dépenses, etc.). Cependant, le consensus implique de déterminer l'ordre des événements dans la blockchain - et de parvenir à un accord sur cet ordre, (effectué par des mineurs)[13].

H. **Chaînage des blocs** : Les blocs sont enchaînés à travers chaque bloc contenant le condensé de hachage de l'en-tête du bloc précédent, après le processus de validation, formant ainsi la chaîne de blocs. Si un bloc précédemment publié était modifié, il aurait un hachage différent. À son tour, cela entraînerait également des hachages différents pour tous les blocs suivants, car ils incluent le hachage du bloc précédent. Cela permet de détecter et de rejeter facilement les blocs modifiés[28].

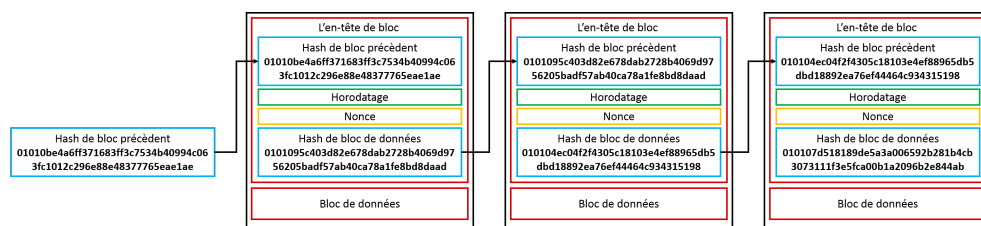


FIGURE 2.9: Chaîne générique de blocs[28]

### 2.3.3 Modèles de consensus

Les modèles de consensus déterminent quel nœud stocker le bloc suivant et comment le nouveau bloc ajouté doit être validé.

A. **Le preuve de travail** : dans le modèle de preuve de travail, Proof of work (PoW), un utilisateur publie le bloc suivant en étant le premier à résoudre un casse-tête intensif en calcul. La solution à ce casse-tête est la «preuve» qu'ils ont effectué le travail. Le casse-tête est conçu de telle sorte qu'il est difficile de résoudre le casse-tête, mais il est facile de vérifier qu'une solution est valide. Cela permet à tous les autres nœuds complets de valider facilement tous les blocs suivants proposés, et tout bloc proposé qui ne satisfait pas le puzzle serait rejeté, utilisé par Bitcoin, Ethereum, et beaucoup d'autres[28].

**Exemples** : dans cet exemple, la chaîne de texte «blockchain» est ajoutée avec une valeur nonce, puis le condensé de hachage est calculé. Les valeurs nonce utilisées seront uniquement des valeurs numériques. Il s'agit d'un casse-tête relativement facile à résoudre et voici quelques exemples de résultats dan la figure 2.10 :

```
SHA256("blockchain0") =
0xbd4824d8ee63fc82392a6441444166d22ed84aaa6dab11d4923075975acab938
(not solved)

SHA256("blockchain1") =
0xdb0b9c1cb5e9c680dffff7482f1a8efad0e786f41b6b89a758fb26d9e223e0a10
(not solved)

...

SHA256("blockchain10730895") =
0x00000ca1415e0bec568f6f605fcc83d18cac7a4e6c219a957c10c6879d67587
(solved)
```

FIGURE 2.10: Exemple de la chaîne de texte «blockchain»[28]

**Objectif** : fournir un obstacle à la publication de blocs sous la forme d'un casse-tête difficile à résoudre pour permettre les transactions entre des participants non fiables.

**Avantages** : ses avantages sont :

- Difficile d'effectuer un déni de service en inondant le réseau de blocs défectueux.
- Ouvert à toute personne disposant de matériel pour résoudre le puzzle.

**Désavantages** : ses désavantages sont :

- Calcul intensif (par conception), consommation d'énergie, course aux armements matériels.
- Potentiel d'attaque de 51% en obtenant suffisamment de puissance de calcul.

B. **Preuve de participation** : Le modèle de preuve de participation, Proof of stake (PoS) est basé sur l'idée que plus l'utilisateur de stake a investi dans le système, plus ils voudront que le système réussisse et moins ils voudront le renverse, utilisé par Ethereum Casper, Krypton[28].

**Objectif :** pour permettre une barrière de calcul moins intensive aux blocs de publication, tout en permettant des transactions entre des participants non fiables.

**Avantages :** ses avantages sont :

- Moins intensif en calcul que PoW.
- Ouvert à tous ceux qui souhaitent miser des crypto-monnaies.
- Les parties prenantes contrôlent le système.

**Désavantages :** ses désavantages sont :

- Les parties prenantes contrôlent le système.
- Rien n'empêche la constitution d'un pool d'acteurs pour créer un pouvoir centralisé.
- Potentiel d'attaque de 51% en obtenant suffisamment de puissance financière.

C. **PoS délégués :** utilisé par Bitshares, Steem, Cardano, EOS[28].

**Objectif :** Permettre un modèle de consensus plus efficace grâce à une «démocratie liquide» où les participants votent (à l'aide de messages signés par cryptographie) pour élire et révoquer les droits des délégués de valider et de sécuriser la blockchain.

**Avantages :** ses avantages sont :

- Les délégués élus sont économiquement incités à rester honnêtes.
- Plus efficace en termes de calcul que PoW.

**Désavantages :** ses désavantages sont :

- Moins de diversité de nœuds que les implémentations de consensus PoW ou PoS pur.
- Risque de sécurité accru pour la compromission des nœuds en raison d'un ensemble restreint de nœuds opérationnels.
- Comme tous les délégués sont «connus», les producteurs de blocs peuvent être incités à s'entendre et à accepter des pots-de-vin, ce qui compromet la sécurité du système.

D. **Round Robin :** est un modèle de consensus utilisé par certains réseaux blockchain autorisés. Dans ce modèle de consensus, les nœuds créent tour à tour des blocs. Round Robin Consensus a une longue histoire ancrée dans l'architecture de systèmes distribués. Pour gérer les situations où un nœud de publication n'est pas disponible pour publier un bloc à son tour, ces systèmes peuvent inclure une limite de temps pour permettre aux nœuds disponibles de publier des blocs afin que les nœuds non disponibles n'entraînent pas un arrêt de la publication des blocs. Ce modèle garantit qu'aucun nœud ne crée la majorité des blocs. Il bénéficie d'une approche simple, manque d'énigmes cryptographiques et nécessite peu d'énergie, utilisé par MultiChain[28].

**Objectif :** fournir un système de publication de blocs parmi les nœuds de publication approuvés.

**Avantages :** ses désavantages sont :

- Faible puissance de calcul.
- Simple à comprendre.

**Désavantage :** nécessite une grande confiance entre les nœuds de publication.

- E. **Preuve d'autorité / d'identité :** le modèle de consensus de preuve d'autorité (également appelé preuve d'identité) repose sur la confiance partielle des nœuds de publication via leur lien connu avec les identités du monde réel. Les nœuds de publication doivent avoir leur identité prouvée et vérifiable dans le réseau de la blockchain (par exemple, les documents d'identification qui ont été vérifiés et notariés et inclus dans la blockchain). L'idée est que le nœud de publication mise son identité / réputation pour publier de nouveaux blocs. Les utilisateurs du réseau Blockchain affectent directement la réputation d'un nœud de publication en fonction du comportement du nœud de publication. Les nœuds de publication peuvent perdre de la réputation en agissant d'une manière avec laquelle les utilisateurs du réseau blockchain ne sont pas d'accord, tout comme ils peuvent gagner en réputation en agissant d'une manière avec laquelle les utilisateurs du réseau blockchain sont d'accord, utilisé par Ethereum Kovan testnet, POA Chain, divers systèmes autorisés utilisant Parity[28].

**Objectif :** créer un processus de consensus centralisé pour minimiser le taux de création et de confirmation de bloc

**Avantages :** ses avantages sont :

- Temps de confirmation rapide.
- Permet des taux de production de blocs dynamiques.
- Peut être utilisé dans les chaînes latérales pour les réseaux de chaînes de blocs qui utilisent un autre modèle de consensus.

**Désavantages :** ses désavantages sont :

- Se fonde sur l'hypothèse que le nœud de validation actuel n'a pas été compromis.
- Conduit à des points de défaillance centralisés.
- La réputation d'un nœud donné est sujette à un risque de queue (tail risk), car elle peut être compromise à tout moment.

- F. **Preuve du temps écoulé :** dans le cadre du modèle de consensus de preuve de temps écoulé, Proof of Elapsed Time (PoET), chaque nœud de publication demande un temps d'attente à une source de temps

matériel sécurisé au sein de son système informatique. La source de temps du matériel sécurisé générera un temps d'attente aléatoire et le renverra au logiciel du nœud de publication. Les nœuds de publication prennent le temps aléatoire qui leur est donné et deviennent inactifs pendant cette durée. Une fois qu'un nœud de publication se réveille de l'état inactif, il crée et publie un bloc sur le réseau de la chaîne de blocs, alertant les autres nœuds du nouveau bloc ; tout nœud de publication qui est encore inactif cessera d'attendre et tout le processus recommence, utilisé par Hyperledger Sawtooth[28].

**Objectif :** permettre un modèle de consensus plus économique pour les réseaux de chaînes de blocs, au détriment de garanties de sécurité plus approfondies associées à PoW.

**Avantage :** moins coûteux en calcul que PoW

**Désavantages :** ses désavantages sont :

- Configuration matérielle requise pour gagner du temps.
- Suppose que l'horloge matérielle utilisée pour dériver l'heure n'est pas compromise.
- Compte tenu des limites de vitesse de latence tardive, la synchronicité en temps réel est essentiellement impossible dans les systèmes distribués.

### 2.3.4 Contrats intelligents

Les contrats intelligents peuvent être considérés comme une grande avancée dans la technologie de la blockchain. Dans les années 1990, un contrat intelligent a été proposé comme protocole de transaction informatisé qui exécute les termes contractuels d'un accord. Les clauses contractuelles qui sont intégrées dans les contrats intelligents seront automatiquement appliquées lorsqu'une certaine condition est remplie (par exemple, une partie qui viole le contrat sera punie automatiquement).

Les blockchains permettent des contrats intelligents. Les contrats intelligents sont essentiellement mis en œuvre au-dessus des chaînes de blocs. Les clauses contractuelles approuvées sont converties en programmes informatiques exécutables. Les connexions logiques entre les clauses contractuelles ont également été préservées sous la forme de flux logiques dans les programmes (par exemple, l'instruction if-else-if). L'exécution de chaque déclaration de contrat est enregistrée comme une transaction immuable stockée dans la blockchain. Les contrats intelligents garantissent un contrôle d'accès et une application des contrats appropriés.

En particulier, les développeurs peuvent attribuer une autorisation d'accès à chaque fonction du contrat. Une fois qu'une condition d'un contrat intelligent est satisfaite, l'instruction déclenchée exécutera automatiquement la fonction correspondante de manière prévisible. Par exemple, Alice et Bob

s'entendent sur la sanction de la violation du contrat. Si Bob rompt le contrat, la pénalité correspondante (comme spécifié dans le contrat) sera automatiquement payée (déduite) du dépôt de Bob.

Le cycle de vie complet des contrats intelligents se compose de quatre phases consécutives (création, déploiement, exécution et achèvement), comme l'illustre la figure 2.11[29] :

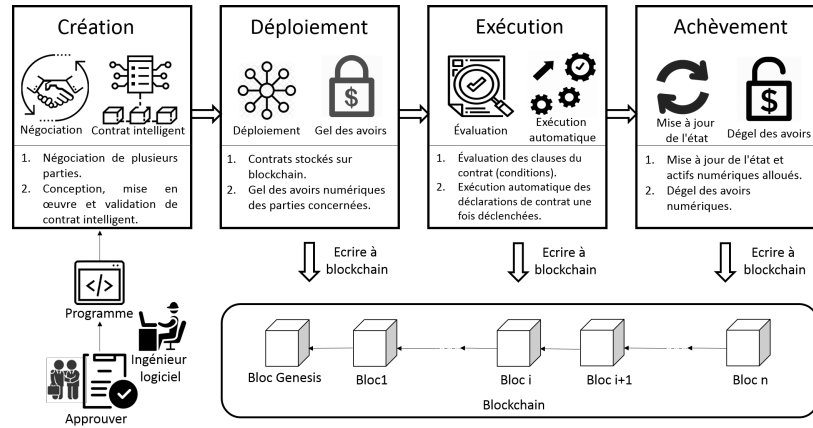


FIGURE 2.11: Un cycle de vie d'un contrat intelligent se compose de quatre phases principales[29]

- A. **Création de contrats intelligents** : Plusieurs parties concernées ont d'abord négocié les obligations, les droits et les interdictions sur les contrats. Après plusieurs cycles de discussions et de négociations, un accord peut être trouvé. Des avocats ou des conseillers aideront les parties à rédiger un premier accord contractuel. Les ingénieurs logiciels convertissent ensuite cet accord écrit en langues naturelles en un contrat intelligent écrit en langages informatiques, y compris les langages déclaratifs et les langages de règles logiques. Semblable au développement de logiciels informatiques, la procédure de conversion du contrat intelligent est composée de la conception, de la mise en œuvre et de la validation (c'est-à-dire des tests). Il convient de mentionner que la création de contrats intelligents est un processus itératif impliquant plusieurs cycles de négociations et d'itérations. Parallèlement, il est également impliqué avec plusieurs parties, telles que les parties prenantes, les avocats et les ingénieurs logiciels.
- B. **Déploiement de contrats intelligents** : Les contrats intelligents validés peuvent ensuite être déployés sur des plateformes au-dessus des chaînes de blocs. Les contrats stockés sur les blockchains ne peuvent pas être modifiés en raison de l'immutabilité des blockchains. Toute modification nécessite la création d'un nouveau contrat. Une fois les contrats intelligents déployés sur les blockchains, toutes les parties

peuvent accéder aux contrats via les blockchains. De plus, les avoirs numériques des deux parties impliquées dans le contrat intelligent sont verrouillés via le gel des portefeuilles numériques correspondants. Par exemple, les transferts de pièces (entrants ou sortants) sur les portefeuilles concernés par le contrat sont bloqués. Pendant ce temps, les parties peuvent être identifiées par leurs portefeuilles numériques.

- C. **Exécution de contrats intelligents** : Après le déploiement des contrats intelligents, les clauses contractuelles ont été suivies et évaluées. Une fois les conditions contractuelles atteintes (par exemple, réception du produit), les procédures contractuelles (ou fonctions) seront automatiquement exécutées. Il convient de noter qu'un contrat intelligent se compose d'un certain nombre d'instructions déclaratives avec des connexions logiques. Lorsqu'une condition est déclenchée, l'instruction correspondante sera automatiquement exécutée, par conséquent une transaction sera exécutée et validée par les mineurs dans les blockchains. Les transactions validées et les états mis à jour ont ensuite été stockés sur les blockchains.
- D. **Achèvement de contrats intelligents** : Après l'exécution d'un contrat intelligent, les nouveaux états de toutes les parties concernées sont mis à jour. En conséquence, les transactions lors de l'exécution des contrats intelligents ainsi que les états mis à jour sont stockés dans des blockchains. Pendant ce temps, les actifs numériques ont été transférés d'une partie à une autre (par exemple, le transfert d'argent de l'acheteur au fournisseur). Par conséquent, les avoirs numériques des parties concernées ont été débloqués. Le contrat intelligent a alors achevé tout le cycle de vie.

## 2.4 Blockchain et la préservation des données

La technologie Blockchain préserve les données en préservant la sécurité et la confidentialité.

- A. **Blockchain au service de sécurité** : la sécurité dans la blockchain peut être définie comme la protection des informations et données de transaction dans un bloc (quelle que soit la forme de données) contre les menaces internes et périphériques, malveillantes et involontaires. Généralement, cette protection implique la détection de la menace, la prévention de la menace, une réponse appropriée à la menace à l'aide de politiques de sécurité, d'outils et de services informatiques, ci-dessous quelques idées et principes importants en matière de sécurité[19] :

**Défense en pénétration**, il s'agit d'une stratégie qui utilise de nombreuses mesures correctives pour protéger les données. Il suit le principe selon lequel la protection des données sur plusieurs couches est plus efficace que sur une seule couche de sécurité.

**Privilège minimum**, dans cette stratégie, l'accès aux données est réduit au niveau le plus bas possible pour renforcer le niveau élevé de sécurité.

**Gérez les vulnérabilités**, dans cette stratégie, nous vérifions les vulnérabilités et les gérons en identifiant, authentifiant, modifiant et corrigeant.

**Gérez les risques**, dans cette stratégie, nous traitons les risques dans un environnement en identifiant, évaluant et contrôlant les risques.

**Gérez les correctifs**, dans cette stratégie, nous corrigeons la partie défectueuse comme le code, l'application, le système d'exploitation, le firmware, etc. en acquérant, testant et installant des correctifs.

- B. **Blockchain au service de confidentialité** la vie privée est la capacité d'une seule personne ou d'un groupe de s'isoler ou d'exprimer des données avec discernement. La vie privée dans la blockchain signifie pouvoir effectuer des transactions sans divulguer d'informations d'identification. Dans le même temps, la vie privée permet à un utilisateur de rester conforme en se dévoilant avec discernement sans présenter son activité à l'ensemble du réseau. L'objectif de l'amélioration de la vie privée dans les chaînes de blocs est de rendre extrêmement difficile pour les autres utilisateurs de copier ou d'utiliser le profil cryptographique d'autres utilisateurs. Un volume incommensurable de variations peut être perçu lors de l'application de la technologie blockchain. Certaines caractéristiques communes sont particulièrement importantes et se résument comme suit[19] :

**Tri des données stockées**, la blockchain offre la flexibilité de stocker toutes les formes de données. La perspective de vie privée dans la blockchain varie pour les données personnelles et organisationnelles. Bien que les règles de vie privée s'appliquent aux données personnelles, des règles de vie privée plus strictes s'appliquent aux données sensibles et organisationnelles.

**Distribution de stockage**, les nœuds du réseau qui stockent des copies complètes de la blockchain sont appelés nœuds complets. Les nœuds complets en combinaison avec la caractéristique d'ajout uniquement de la blockchain entraînent une redondance des données. Cette redondance des données prend en charge deux caractéristiques clés de la technologie blockchain, notamment la transparence et la vérifiabilité. La compatibilité de l'application avec la minimisation des données décide du niveau de transparence et de vérifiabilité de ce réseau pour une application.

**Ajouter uniquement**, il est impossible de modifier les données des blocs précédents dans la blockchain sans être détecté. La fonctionnalité d'ajout uniquement de la blockchain dans certains cas



ne restreint pas le droit à la correction des utilisateurs, surtout si les données sont enregistrées de manière incorrecte. Une attention particulière doit être accordée lors de l'attribution des droits aux personnes concernées dans la technologie de la blockchain.

**Blockchain privée vs publique,** l'accessibilité de la blockchain est remarquable du point de vue de la vie privée. À un niveau avancé, les données restreintes sur un bloc peuvent être chiffrées pour un accès conditionnel par les utilisateurs autorisés, car chaque nœud de la chaîne de blocs possède une copie de l'ensemble de la blockchain.

**Types de blockchain non autorisés vs autorisés,** avec les applications de blockchain publiques ou non autorisées, tous les utilisateurs sont en principe autorisés à ajouter des données. Permettre la restauration de médiateurs de confiance influence la répartition du contrôle sur le réseau.

Les soins médicaux sont devenus une partie indispensable de la vie des gens, avec une augmentation spectaculaire du volume de données médicales (par exemple, les certificats de diagnostic et les dossiers médicaux). Cependant, les données médicales sont facilement volées, falsifiées ou même complètement supprimées. Si ce qui précède se produit, les données médicales ne peuvent pas être enregistrées ou récupérées de manière fiable, ce qui retarde la progression du traitement, voire met en danger la vie du patient[22].

Plusieurs systèmes ont été proposés pour la conservation des données lors de l'application de la technologie blockchain. nous définissons certains travaux dans les sous-sections suivantes.

#### 2.4.1 Système de conservation des données basé sur la blockchain pour les données médicales

Dans ce travail, ils ont proposé un nouveau système de conservation des données basé sur la blockchain (DPS) pour les données médicales. Pour fournir une solution de stockage fiable afin d'assurer la primitivité et la vérifiabilité des données stockées tout en préservant la confidentialité des utilisateurs, ils ont exploité le cadre de la blockchain. Avec le DPS proposé, les utilisateurs peuvent conserver des données importantes à perpétuité et l'originalité des données peut être vérifiée si une falsification est suspectée. En outre, ils ont utilisé des stratégies de stockage de données prudentes et une variété d'algorithmes cryptographiques pour garantir la confidentialité des utilisateurs; Par exemple, un adversaire est incapable de lire le texte brut même si les données sont volées. ils ont implémenté un prototype du DPS basé sur la plate-forme du monde réel qui basée sur la blockchain Ethereum. Les résultats de l'évaluation du rendement démontrent l'efficacité et l'efficience du système proposé[22].

**L'idée de système :** leur idée de base est d'introduire le concept de preuve de primitivité des données et d'utiliser la structure de données unique et la décentralisation de la blockchain pour développer un système de conservation des données[22].

**Cadre du système :** la figure 2.12 montre l'opération de conservation des données[22].

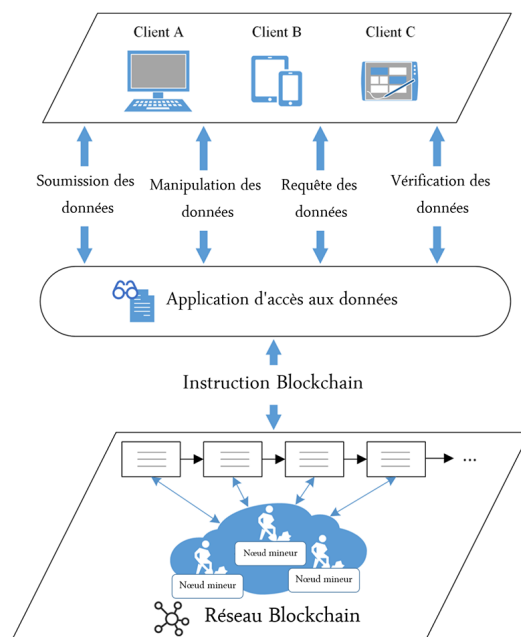


FIGURE 2.12: L'opération de conservation des données[22]

**Avantages,** parmi les avantages on trouve :

- Ce système peut résister efficacement aux opérations de falsification ou de suppression de données, et détecter les opérations illégales des données et informer les utilisateurs à temps.
- L'anonymat du système.
- Il peut gérer des situations dans lesquelles les données sont facilement perdues et altérées.

**Limitations,** parmi les limites de ce système :

- Utilisation d'Ethereum une blockchain publique qui n'a pas de détails spécifiques sur les personnes participant au réseau, il y a donc un risque de collecter les données.
- L'utilisation d'applications de portefeuille peut être menacée par des attaques ou la perte de clé.



### 2.4.3 Une blockchain décentralisée de soins de santé préservant la confidentialité pour l'IoT

Dans ce travail, ils ont tenté de résoudre les problèmes mentionnés ci-dessus liés à l'utilisation de la blockchain avec des appareils IoT. ils ont proposé un nouveau cadre de modèles de chaîne de blocs modifiés adaptés aux appareils IoT qui reposent sur leur nature distribuée et d'autres propriétés de confidentialité et de sécurité supplémentaires du réseau. Ces propriétés de confidentialité et de sécurité supplémentaires dans notre modèle sont basées sur des primitives cryptographiques avancées. Les solutions présentées rendent les données et les transactions des applications IoT plus sécurisées et anonymes sur un réseau basé sur la blockchain[14].

**L'idée de système :** un premier aperçu d'un modèle IoT basé sur la blockchain entrevu dans un modèle avancé de sécurité et de confidentialité à utiliser dans tout système de surveillance à distance IoT actuel[14].

**Réseau de superposition de système :** La figure 2.14 présente le principe de réseau de superposition de système[14].

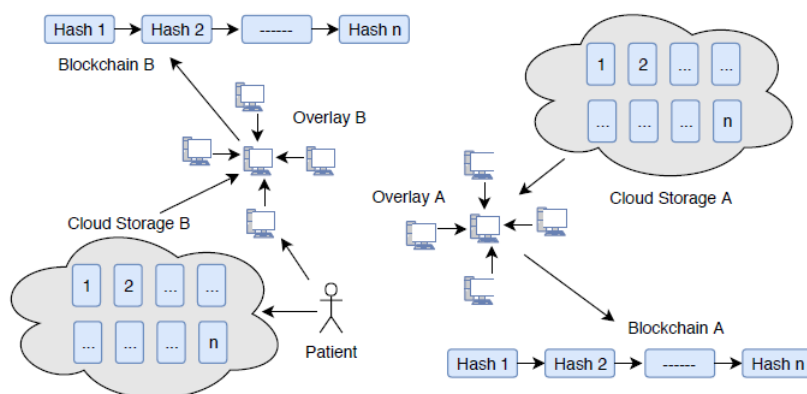


FIGURE 2.14: Réseau de superposition de ce système[14]

**Avantages,** parmi les avantages on trouve :

- Ce système rendent les données et les transactions des applications IoT plus sécurisées et anonymes sur un réseau basé sur la blockchain,
- Ce système est abordé par les concepteurs de modèles : confidentialité, intégrité et disponibilité.

**Limitations,** parmi les limites de ce système :

- Le système ne traite pas la partie transparence et traçabilité entre le patient et les autres parties du système.

#### 2.4.4 Système de Blockchain de Healthcare utilisant des contrats intelligents pour une surveillance à distance automatisée et sécurisée des patients

Dans ce travail, ils ont créé un système où les capteurs communiquent avec un appareil intelligent qui appelle des contrats intelligents et écrit des enregistrements de tous les événements sur la blockchain. Ce système de contrat intelligent prend en charge la surveillance des patients en temps réel et les interventions médicales en envoyant des notifications aux patients et aux professionnels de la santé, tout en conservant un enregistrement sécurisé des personnes qui ont lancé ces activités[16].

**L'idée de système :** ce système résoudrait de nombreuses failles de sécurité associées à la surveillance à distance des patients et automatiserait la remise des notifications à toutes les parties impliquées d'une manière conforme à la HIPAA activités[16].

**La conception de système :** la figure 2.15 montre que les données formatées de l'appareil intelligent sont envoyées au contrat intelligent, qui traite et exécute les actions nécessaires en fonction des résultats et des paramètres prédéterminés activités[16] :

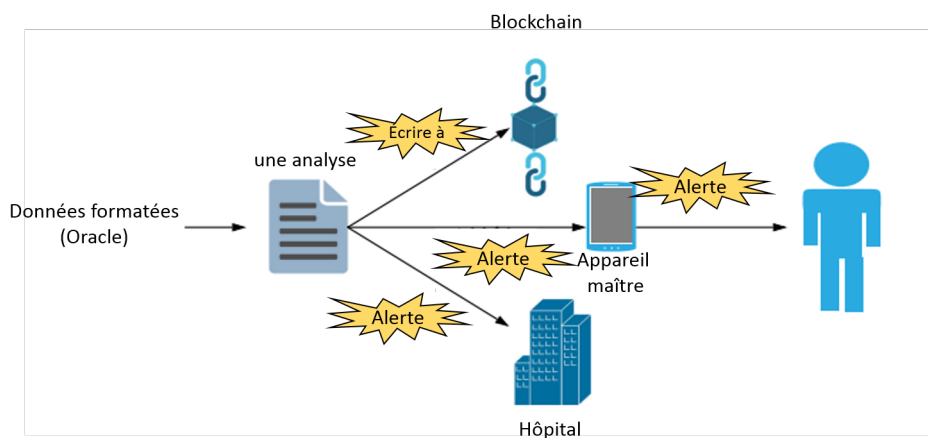


FIGURE 2.15: Conception de système[16]

**Avantages,** Le système respecte la confidentialité, la disponibilité, l'immuabilité, la traçabilité, la vitesse, la confidentialité et la transparence.

**Limitations,** les limites de système sont :

- Le défi de maintenir la sécurité à chaque nœud individuel.
- La gestion des clés peut devenir un problème lorsqu'il existe de nombreux appareils intelligents diffusant leurs transactions vers plusieurs nœuds en attente de vérification du bloc suivant.

- Il doit y avoir un nombre suffisant de nœuds en ligne à tout moment afin de répondre aux exigences pour fournir le nombre minimum de signatures de validation et maintenir l'intégrité de l'algorithme de consensus.

### 2.4.5 Synthèse

Nous avons réalisé une étude approfondie de chaque système en termes d'idée, cadre de système, d'architecture, de réseau ou de conception, ainsi que les avantages et les limitations.

Le tableau 2.3 est une comparaison entre les 4 systèmes, selon la confidentialité, l'autorité, le contrôle utilisateur, l'intégrité et l'anonymat.

| Système 1            |  |
|----------------------|--|
| Confidentialité      | <ul style="list-style-type: none"> <li>• L'utilisation de plusieurs algorithmes de cryptographie.</li> </ul>   |
| Autorisation         | <ul style="list-style-type: none"> <li>• Signature numérique.</li> <li>• L'application portefeuille.</li> </ul>  |
| Contrôle utilisateur | <ul style="list-style-type: none"> <li>• Le système supporte des différents environnements, applications et objet d'accès.</li> <li>• Le programme d'accès aux données : <ul style="list-style-type: none"> <li>— Les utilisateurs soumettent des données à des fins de conservation, qui sont traitées par le système.</li> <li>— Les utilisateurs peuvent interroger les données conservées et vérifier la primitivité.</li> </ul> </li> </ul> |
| Intégrité            | <ul style="list-style-type: none"> <li>• Fonction SHA-256.</li> <li>• Les deux types de contrats utilisés qui assurent l'intégrité.</li> </ul>   |
| Anonymat             | <ul style="list-style-type: none"> <li>• Les données chiffrés.</li> </ul>  |

| Système 2            |   |
|----------------------|---|
| Confidentialité      | <ul style="list-style-type: none"> <li>• Un réseau Ethereum privé.</li> </ul>   |
| Autorisation         | <ul style="list-style-type: none"> <li>• DNS-identity management offre une expérience de type système de noms de domaine pour la gestion des identités d'attributs vérifiés.</li> <li>• Contrats intelligents.</li> <li>• Registre Ethereum autorisé : les données doivent être stockées de manière à rendre impossible pour les parties indésirables d'y accéder.</li> <li>• Algorithme de signature numérique (génération de clés Ethereum).</li> </ul> |
| Contrôle utilisateur | <ul style="list-style-type: none"> <li>• Les utilisateurs ont le droit d'écrire et de lire les données.</li> <li>• Les autres parties peuvent lire les données mais avec une divulgation minimale.</li> </ul>   |
| Intégrité            | <ul style="list-style-type: none"> <li>• Hachage des données.</li> </ul>  |
| Anonymat             | <ul style="list-style-type: none"> <li>• Une identité numérique pour les données.</li> </ul>  |
| Système 3            |   |
| Confidentialité      | <ul style="list-style-type: none"> <li>• Cryptographie symétrique et asymétrique.</li> <li>• Preuve d'autorité.</li> </ul>  |
| Autorisation         | <ul style="list-style-type: none"> <li>• Clé public Digital.</li> <li>• Signature numérique.</li> </ul>   |

|                      |  |
|----------------------|--|
| Contrôle utilisateur | <ul style="list-style-type: none"> <li>• Preuve d'autorité.</li> </ul>   |
| Intégrité            | <ul style="list-style-type: none"> <li>• Le hachage des blocs de données.</li> <li>• Merkle Tree.</li> </ul>   |
| Anonymat             | <ul style="list-style-type: none"> <li>• Les données chiffrés.</li> </ul>  |
| Système 4            |  |
| Confidentialité      | <ul style="list-style-type: none"> <li>• pre-authorized (mining).</li> </ul>   |
| Autorisation         | <ul style="list-style-type: none"> <li>• Les transactions de la blockchain peuvent être retracées depuis l'origine de la création avec une immuabilité garantie et sont signées par les vérificateurs.</li> </ul>        |
| Contrôle utilisateur | <ul style="list-style-type: none"> <li>• Les patients sont capables de relier les actions de surveillance à distance directement à leur dossier médical tout en préservant la confidentialité et le contrôle.</li> </ul> |
| Intégrité            | <ul style="list-style-type: none"> <li>• Les blocs vérifiés sont immuables et résilients à tous les types de manipulation.</li> </ul>  |
| Anonymat             | <ul style="list-style-type: none"> <li>• Les adresses anonymes protégeront l'identité des patients, par conséquent aucune association ne peut être établie entre les patients et leurs données.</li> </ul>               |

TABLE 2.3: Comparaison des systèmes



## 2.5 Conclusion

Dans ce qui précède, nous avons présenté la sécurité et la blockchain, nous avons vu certains systèmes utilisent la Blockchain pour la conservation des données personnelles, leurs avantages et leurs limites ainsi qu'une comparaison selon la confidentialité, l'autorisation, le contrôle des utilisateurs, l'intégrité et l'anonymat. Dans le chapitre suivant, nous présenterons notre contribution basée sur des nouvelles techniques et une nouvelle architecture de blockchain pour améliorer la sécurité des systèmes de diagnostic médical en ligne.

# Chapitre 3

**Contribution : une architecture basée blockchaine et certificat intelligent pour protéger les données personnelles.**

---

*"La blockchain aura un impact énorme sur le secteur de la technologie, la saluant comme... la prochaine grande révolution informatique qui est sur le point de se produire."*

*Steve Wazniak*

---

## 3.1 Introduction

Nous avons vu que la blockchain est devenue une technologie que nous pouvons utiliser avec ses différents composants pour le stockage des transactions, dans le but de préserver la confidentialité et d'assurer la sécurité des systèmes en ligne.

Dans ce chapitre, nous précisons notre contribution en définissant notre système Healthcare et le système cloud ainsi que le problème de confidentialité et notre solution proposée, les deux niveaux de sécurité et enfin la conception de notre système.

## 3.2 Système Healthcare

Les systèmes de soins de santé, Healthcare, sont les systèmes par lesquels le secteur de la santé peut être réglementé en termes de services et

de dépenses fournis à la population, et avec l'intervention de la technologie au cours des deux derniers siècles pour améliorer la qualité et faciliter les services de santé, ce système est devenu une menace de pénétrer la confidentialité des données des patients en les collectant ou en révélant l'identité des patients et leurs informations personnelles, ce qui peut causer des dommages psychologiques, contre la loi et le droit de la personne à préserver sa vie privée.

Pour améliorer la sécurité des données et garantir qu'elles ne sont pas collectées, nous suggérons le système de soins de santé illustré dans la figure 3.1, où toutes les personnes liées au système de soins de santé communiquent entre elles via un ensemble de blockchains qui garantissent l'intégrité des données et assurent le cryptage et la fragmentation des données sur plusieurs bases de données pour garantir qu'elles ne soit pas accessibles facilement.

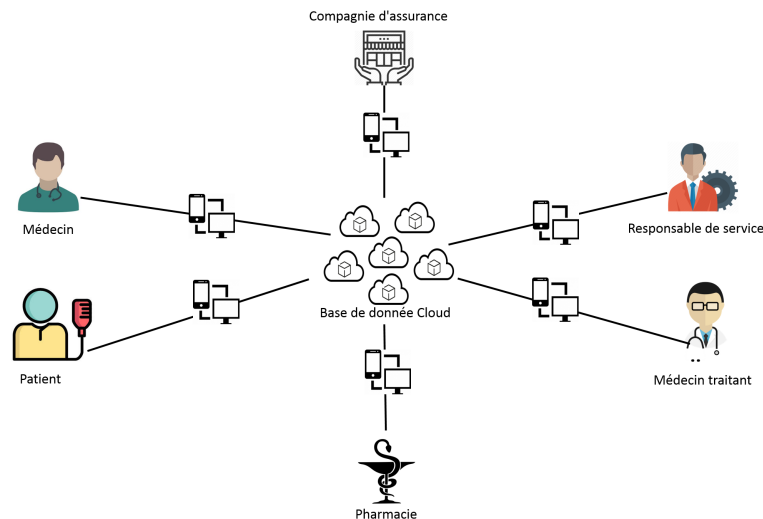


FIGURE 3.1: Système Healthcare

### 3.3 Conception de notre système

La figure 3.2 présente le diagramme de contexte de notre système pour montrer les principales missions des acteurs. Nous avons également défini un diagramme d'activité et un diagramme d'état-transitions pour expliquer le processus de la création de notre système.

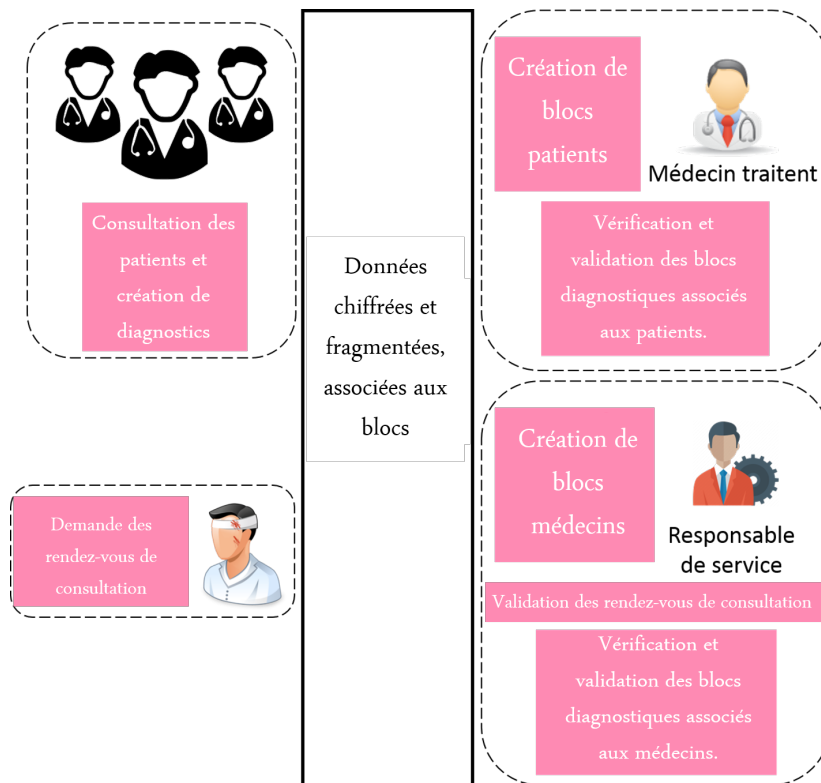


FIGURE 3.2: Diagramme de contexte de notre système

### 3.3.1 Diagramme d'activité

Nous décrivons le flux de travail avec le diagramme d'activité montré dans la figure 3.3 qui aborde le système d'un point de vue globale.

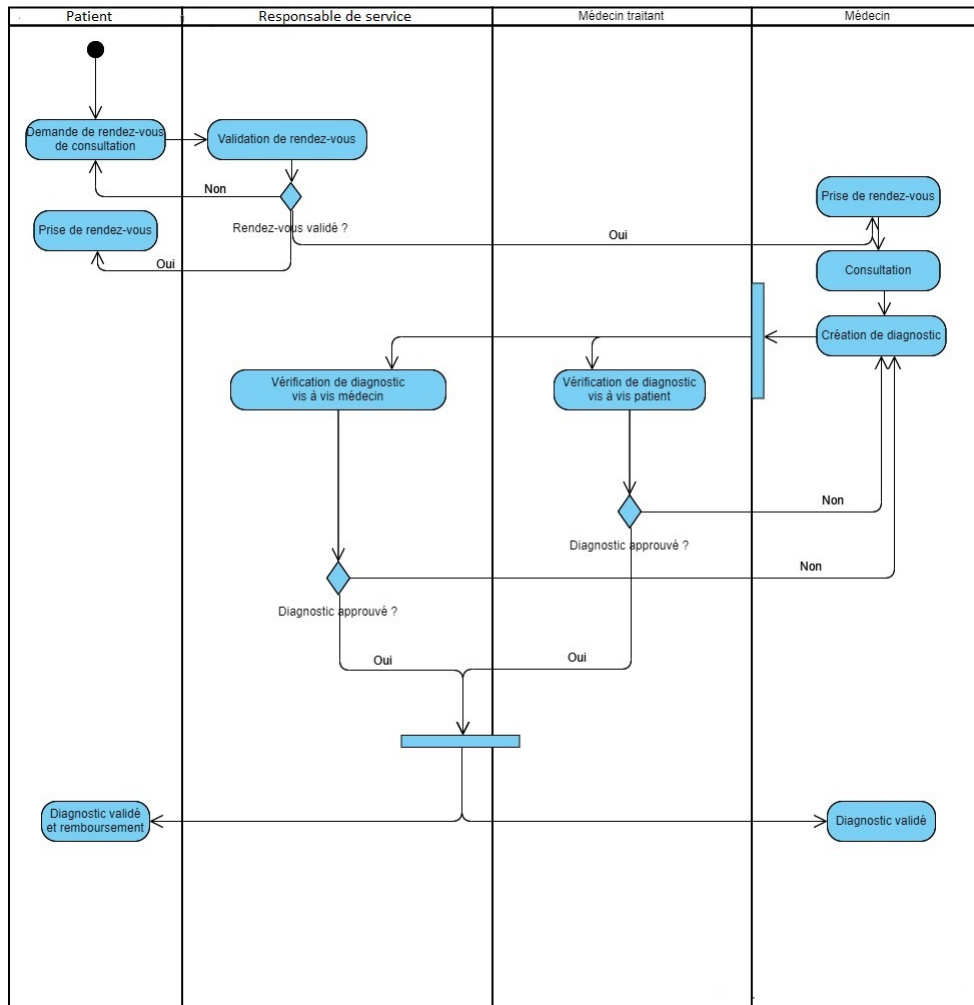


FIGURE 3.3: Diagramme d'activité

### 3.3.2 Diagramme d'états-transitions

le diagramme des états de transition illustré sur la figure 3.4 cible pour présenter les états et les transitions qui sont exécutés en parallèle dans le système de l'acteur patient.

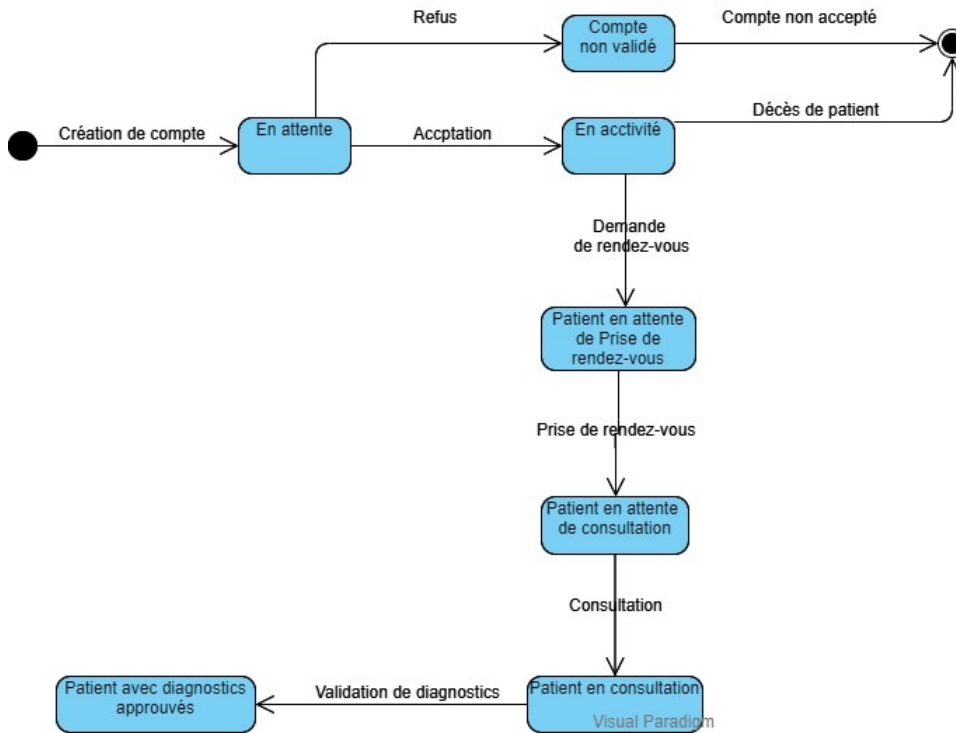


FIGURE 3.4: Diagramme d'états-transitions

### 3.4 Système cloud

Les systèmes cloud jouent un rôle important dans les soins de santé, car ils accélèrent la communication et assurent la surveillance des patients à distance, ainsi que l'augmentation des données médicales, ils peuvent faciliter le stockage, mais les problèmes de sécurité dans les systèmes cloud sont inévitables, d'autant plus que nous ne connaissons pas l'emplacement de stockage et ne faisons pas confiance à ces infrastructures clouds, n'importe qui peut accéder aux données et divulguer son identité, ou les utiliser à des fins commerciales ou personnelles, de sorte que les systèmes cloud sont devenus une partie à double tranchant, la première est positive car elle peut faciliter de nombreux problèmes de stockage, tandis que la seconde est négative en termes de vie privée et de confidentialité des données et leur sécurité, et tout le monde se bat pour accroître la sécurité de ces systèmes, en particulier ceux utilisés dans les soins de santé car ce sont des données sensibles et ne doivent pas être utilisées, ou identifiées uniquement par les personnes concernées.

### 3.4.1 Problème de confidentialité

Nous traitons les problèmes de confidentialité des patients lors de l'utilisation du système de diagnostic sur des plateformes en ligne, où les services sont disponibles pour tous les internautes. Ces applications collectent en permanence des données personnelles à haute résolution, dont l'utilisateur n'a aucune connaissance ni contrôle spécifiques. Notant que le même système pourrait être utilisé pour d'autres problèmes, tels que le partage de données médicales pour la recherche scientifique et des fins commerciales. D'autre part, la vie privée du patient est menacé par différents risques d'intrusion surtout de la part des pirates. À la lumière de cela, notre système protège contre certains problèmes de confidentialité :

**L'anonymat des données :** nous cherchons à cacher l'identité, et à la montrer uniquement aux personnes autorisées, dans le but de garantir l'anonymisation des données personnelles des patients et des médecins, même si quelqu'un accède ou collecte les données, celles-ci ne peuvent pas être démontées ni traitées.

**Transparence :** chaque patient dispose d'une totale transparence sur ses données et la manière dont elles sont accessibles, tant qu'il existe un contrat entre eux et les médecins, validé par les tiers de confiance qui les choisit.

**Contrôle d'accès :** les médecins et les patients ne peuvent accéder qu'à leurs données, le médecin ne peut accéder qu'à ses patients, et les tiers valident toutes les transactions, à savoir : l'inscription, les rendez-vous de consultation et le contrat de diagnostic conclu entre le patient et le médecin qui est résilié après un délai déterminé. Ainsi que chaque valideur ne peut accéder qu'à son patient ou médecin qui les choisit.

**Confidentialité :** seules les personnes autorisées peuvent accéder aux informations limitées à leurs besoins dans le système.

**Traçabilité :** en général, la technologie Blockchain permet la traçabilité, de plus notre système garantit que le diagnostic est lié au patient et au médecin spécifique, et que la trace est conservée.

**Disponibilité :** nous utilisons la technique de la fragmentation par la division de toutes les données en petits morceaux dans différentes bases de données, permettant de maintenir le bon fonctionnement du système, et leur disponibilité.

**Non-répudiation :** seulement des données valides et utiles seront disponibles pour l'utilisateur réel.

**Intégrité :** grâce aux fonctions de hachage, de segmentation et de duplication des données, nous garantissons l'intégrité des données.

### 3.4.2 Solutions suggérées

Nous présentons un aperçu du système proposé, notre idée de base est de prouver l'utilisation de la technologie blockchain d'une manière lisible et réussie, nous l'utilisons sur une base de données cloud, avec la fragmentation et la cryptographie pour assurer et améliorer la sécurité des données, nous lions un contrat intelligent pour limiter l'accès aux données après une période déterminée, et pour effectuer certaines transactions de manière authentique et automatique, afin de développer un système de conservation des données.

#### La fragmentation des donnée

La fragmentation est un concept parmi les conceptions de distribué, l'objectif principal de la fragmentation des données est de faciliter l'accès des utilisateurs aux données, car elles sont réparties en fonction de l'emplacement d'utilisateur, elle permette également la fiabilité, ce qui augmente l'efficacité des requêtes en réduisant la taille de la table à un sous-ensemble plus petit et en lui fournissant moins de latence du réseau. Également parmi les principales caractéristiques de la fragmentation des données, on trouve la sécurité et la confidentialité des données, car elles ne peuvent pas être facilement collectées ou impossibles à collecter, les données sont divisées sur plusieurs bases de données et stockées dans la même base de données avec d'autres données non liées les unes aux autres, ce qu'on appelle la fragmentation mixte, comme le montre la figure 3.5.

| BDD1   | BDD2   | BDD3   | BDD4   | BDD5   |
|--|--|--|--|--|
| <ul style="list-style-type: none"> <li>•Bloc(a)</li> <li>•Donnée1</li> <li>•Bloc(d)</li> <li>•Donnée3</li> </ul> | <ul style="list-style-type: none"> <li>•Bloc(c)</li> <li>•Donnée1</li> </ul>                                     | <ul style="list-style-type: none"> <li>•Bloc(b)</li> <li>•Donnée1</li> <li>•Bloc(a)</li> <li>•Donnée3</li> </ul> | <ul style="list-style-type: none"> <li>•Bloc(a)</li> <li>•Donnée2</li> <li>•Bloc(c)</li> <li>•Donnée2</li> </ul> | <ul style="list-style-type: none"> <li>•Bloc(d)</li> <li>•Donnée2</li> <li>•Bloc(c)</li> <li>•Donnée3</li> </ul> |
| <ul style="list-style-type: none"> <li>•Bloc(b)</li> <li>•Donnée2</li> </ul>                                     | <ul style="list-style-type: none"> <li>•Bloc(f)</li> <li>•Donnée1</li> <li>•Bloc(a)</li> <li>•Donnée4</li> </ul> | <ul style="list-style-type: none"> <li>•Bloc(b)</li> <li>•Donnée3</li> <li>•Bloc(f)</li> <li>•Donnée3</li> </ul> | <ul style="list-style-type: none"> <li>•Bloc(f)</li> <li>•Donnée2</li> </ul>                                     | <ul style="list-style-type: none"> <li>•Bloc(d)</li> <li>•Donnée1</li> </ul>                                     |

FIGURE 3.5: Exemple de la fragmentation mixte

#### Stockage de données hors chaîne

L'utilisation de la blockchain dans la sécurité des données volumineuses crée un problème car nous ne savons pas où iront nos données ni comment elles seront stockées, qu'elles soient centralisées ou décentralisées, ou pour qui nous paierons pour le stockage de nos informations dans la blockchain et si les tiers payés peuvent être fiables ou non fiables à partir de la méthode



utilisée pour le stockage. Nous utilisons donc d'autres solutions, tandis que les données brutes peuvent être stockées comme nous le souhaitons.

Par exemple, nous pouvons utiliser une base de données ou simplement un système de fichiers. Dans notre système nous utilisons une base de données en temps réel (base de données hébergée dans le cloud), pour enregistrer les données de manière structurée et noSQL, ce stockage regroupe les données dans des blocs identiques associés à un numéro unique comme le montre la figure 3.6.

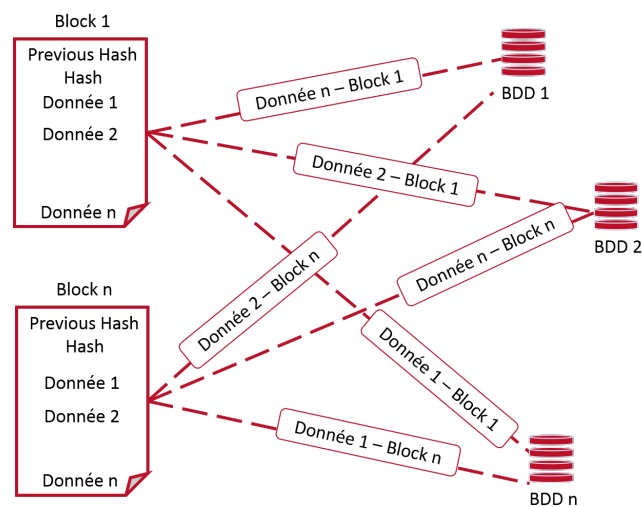


FIGURE 3.6: La fragmentation des données selon le numéro de bloc

Nous nous assurons que nous attribuons l'id (hash) de la transaction blockchain aux données brutes, ainsi ces données de bloc sont associées au hachage de bloc avec un numéro unique comme le montre l'exemple dans la figure 3.5. les données sont fragmentés selon le numéro de bloc a, b, c, d, f.

### Valideurs au lieu de mineurs

Lorsque la blockchain est devenue une open source tout le monde a adopté les mêmes techniques, tandis que le bitcoin utilise des méthodes mathématiques complexes pour le mécanisme de consensus. Beaucoup de gens perçoivent les concepts de la blockchain et de l'exploitation minière comme quelque chose d'indivisible et s'accompagnant les uns des autres, ainsi que de nombreuses fonctionnalités séparent la blockchain Bitcoin d'une blockchain conçue par l'entreprise.

Actuellement, la technologie de la chaîne de blocs peut exister sans extraction, fonctionnant plutôt comme des ordinateurs de réseau cloud, où les nœuds distribués sont fiables et contrôlés. Par conséquent, nous avons utilisé une technique de validation simple dans notre système, tandis que les vali-

deurs sont des tiers de confiance, responsables de la maintenance du nœud sur le réseau et de la validation de toutes les transactions.

- L'administrateur de système valide les nouveaux blocs des médecins traitants et des responsables de service.
- Le médecin traitant valide les blocs des nouveaux patients.
- Le responsable de service valide les blocs des nouveaux médecins ainsi que les rendez-vous de patients.
- Le diagnostic est validé aux deux tiers (chef de service et médecin traitant), alors que le médecin possède l'historique des diagnostics de ses patients, et le responsable du service possède l'historique des diagnostics effectués par ses médecins.

On note que chaque patient choisit son médecin traitant et que chaque médecin choisit son responsable de service.

### Présentation des blocs

Le bloc standard permet la traçabilité, mais dans le cas du diagnostic, il ne suffit pas d'assurer la traçabilité du médecin qui a posé le diagnostic et du patient diagnostiqué, de façon efficace, nous avons donc proposé nouveau modèle de bloc pour les diagnostics, qui ne prend pas seulement le hash et le hash précédent mais aussi le hash du patient diagnostiqué et du médecin qui fait le diagnostic afin d'augmenter le niveau de confidentialité des données à travers l'utilisation des hashes patient et médecin au lieu des information en claire sur ces derniers.

- A. **Blocs standard** : les blocs, médecin traitant, responsable de service, médecin, patient et rendez-vous sont des blocs standard uniquement avec le hash et le hash précédent, comme indiqué dans les exemples des figures 3.7 et 3.8.

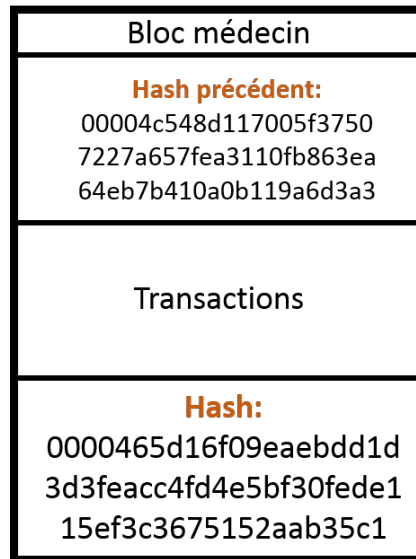


FIGURE 3.7: Bloc d'un médecin

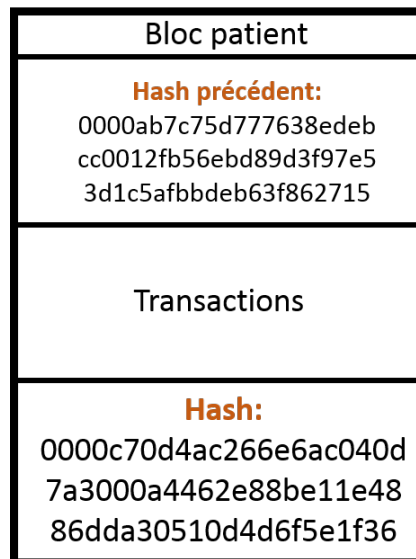


FIGURE 3.8: Bloc d'un patient

B. **Un nouveau modèle de bloc** : la figure 3.9 montre un exemple de bloc diagnostique de notre modèle proposé.

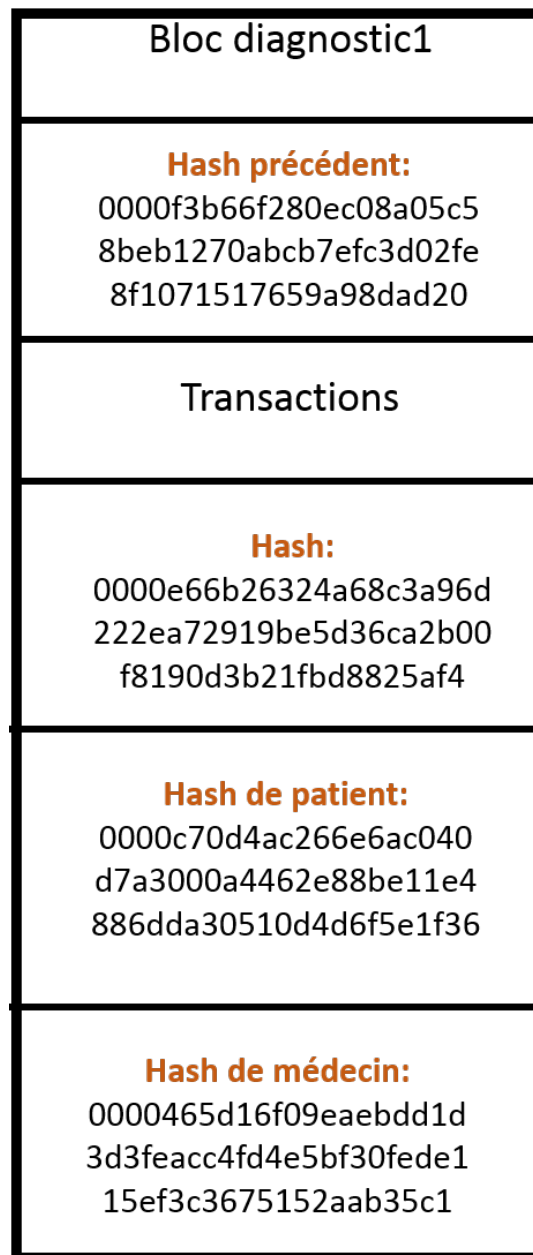


FIGURE 3.9: Bloc d'un diagnostic

### Présentation de l'enchaînement de blocs

Les blocs précédents (médecin traitant, responsable de services, patient, médecin et rendez-vous) sont des blocs standard, donc le séquençage des blocs est de la manière standard.

La figure 3.11 montre la présentation standard de la chaîne de blocs du

patient, et la figure 3.10 montre la séquence des blocs médecin, où chaque bloc est concaténé avec l'autre par le hash et le hash précédent.

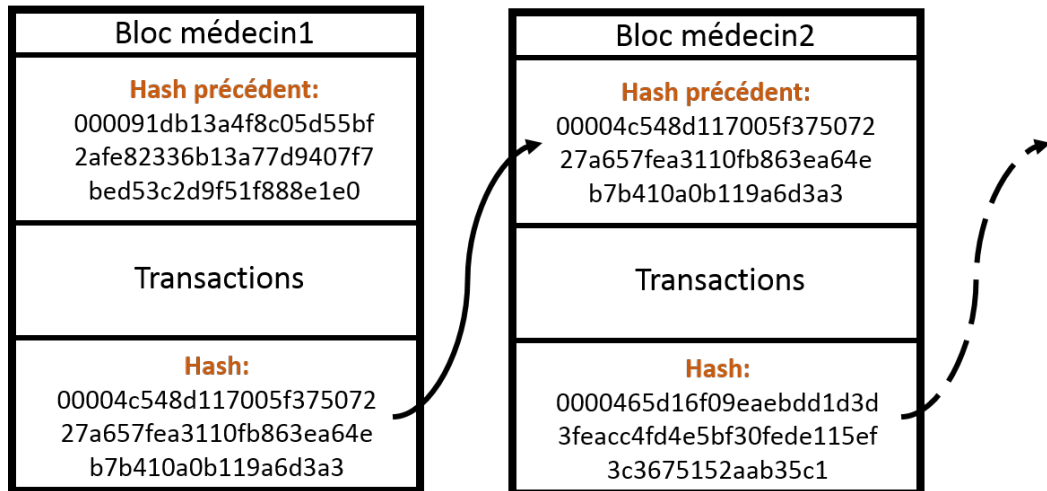


FIGURE 3.10: Exemple d'une chaîne de blocs médecin

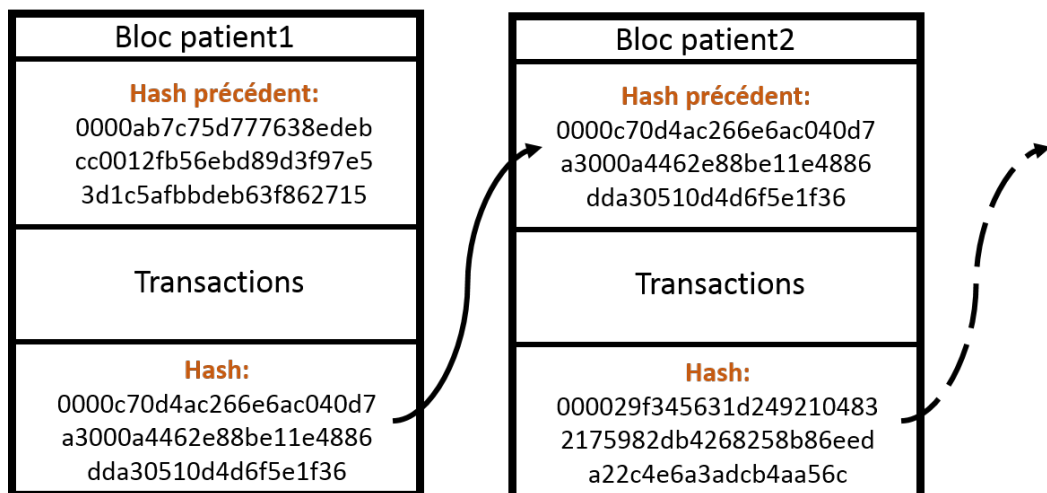


FIGURE 3.11: Exemple d'une chaîne de blocs patient

**Blockchain diagnostic :** le bloc a le hash précédent et le hash, ainsi que le hash du médecin et du patient alors que leur séquence a deux présentations comme indiqué dans la figure 3.12 :

- la présentation standard concatène les blocs de diagnostic avec le code de hash et le code de hash précédent (Bloc diagnostic 1, Bloc diagnostic 2 etc ...).

- la nouvelle présentation qui concatène le diagnostic avec le bloc de patient diagnostiqué par le hash de patient, et avec le bloc de médecin qui fait le diagnostic par le hash de médecin.

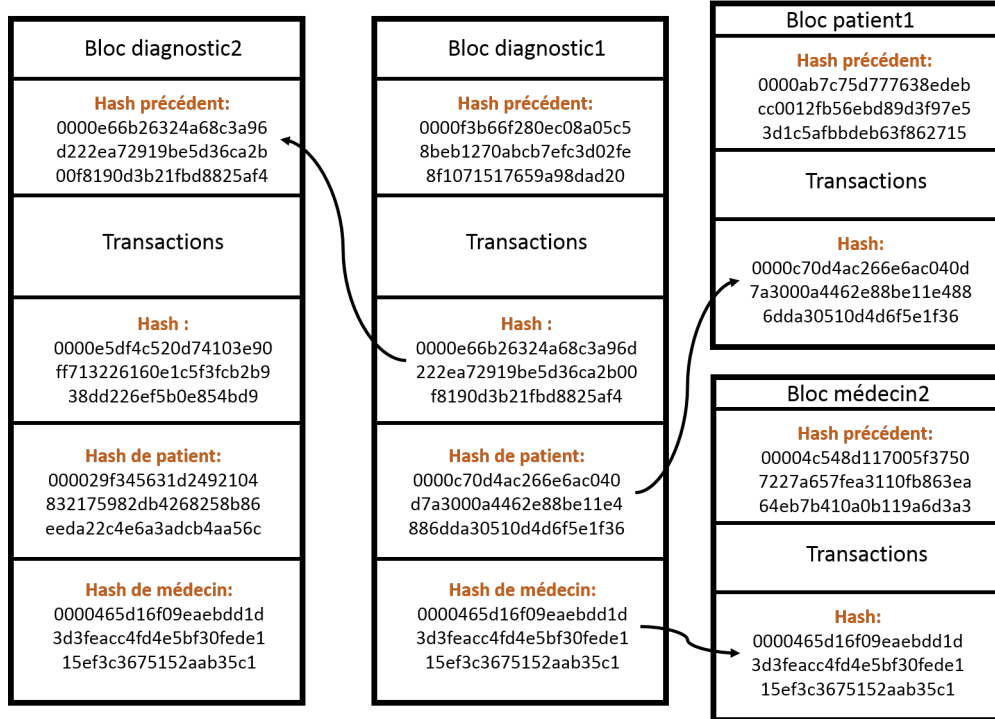


FIGURE 3.12: Exemple d'une chaîne de blocs diagnostique

## Contrats intelligents

Les contrats intelligents sont des protocoles qui sont utilisés automatiquement par le système, pour faciliter la négociation entre les parties au contrat. La technologie de la blockchain a un impact sur l'authentification de ce dernier, alors que les contrats facilitent les transactions de la blockchain, les deux termes sont donc liés l'un à l'autre.

nous intégrons le protocole smart contract dans les cas suivants :

- Nous l'avons utilisé pour déterminer la période de diagnostic par le médecin, puis les tiers de confiance le valident, de sorte que s'il est validé, il se retrouve sur l'interface patient et le médecin qui a posé le diagnostic, si la période prédéterminée est expirée, le diagnostic n'apparaît pas aux deux.
- Remboursement au patient le prix de la consultation ou du médicament, selon le pourcentage de type de maladie.

### 3.4.3 Niveaux de sécurité

La sécurité des données est réalisé à travers l'utilisation du nouveau modèle de block qui est basé sur des imbrication des blockchain. Ce modèle assure une double sécurité le premier niveau à travers le hachage et le cryptage des données du bloc, le deuxième niveau à travers l'utilisation des hash enregistrer dans d'autre blockchain.

**Intégrité des données :** l'intégrité dans notre système est indiquée ci-dessous :

- Maintenir l'intégrité des informations à travers le hashage des données, afin qu'elles ne soient pas pénétrées ou modifiées.
- Duplication en enregistrant les données dans d'autres bases de données.
- Récupération automatiquement des données en cas de modification.

**Confidentialité des données :** la confidentialité dans notre système est indiquée ci-dessous :

- Le chiffrement des données augmente la confidentialité.
- La fragmentation des données réduit le risque de collecte de données.
- Les utilisateurs ne peuvent accéder qu'à la partie qui leur est attribuée.
- Le système s'appuie sur des contrats intelligents qui spécifient la date à laquelle les informations de patient sont affichées ou masquées.
- La nouvelle architecture de bloc diagnostic permet de suivre efficacement les médecins en cas d'erreurs médicales ou d'usurpation d'identité.

## 3.5 Conclusion

Dans ce chapitre, nous avons montré en détail notre système de protection des données des patients, à travers un nouveau modèle pour les blocs, une segmentation des données et l'intégration des contrats intelligents. Le quatrième chapitre présente l'implementation et les scénarios de développement de notre système Healthcare.

# Chapitre 4

## Implémentation et conception

---

*"La blockchain est un tour de force technologique."*

*Bill Gates*

---

### 4.1 Introduction

Nous avons vu ce qui précède, la contribution de notre étude, à savoir principalement les solutions proposées et la conception qui nous aident pour la mise en œuvre, donc dans ce chapitre, nous voyons l'implantation, commencé par les scénarios jusqu'à la description des différentes interfaces et du stockage off-chain.

### 4.2 Scénarios d'étude

Nous montrons comment les acteurs et la plateforme communiquent entre eux en termes de séquence de messages à partir les scénarios suivants.

#### 4.2.1 Diagramme de séquence de création d'un compte médecin traitant ou responsable de service

La figure 4.1 explique la chronologie des messages entre un médecin traitant / responsable de service, la plateforme et l'administrateur en cas d'un compte valide, et l'autre figure 4.2 en cas d'un compte invalide.



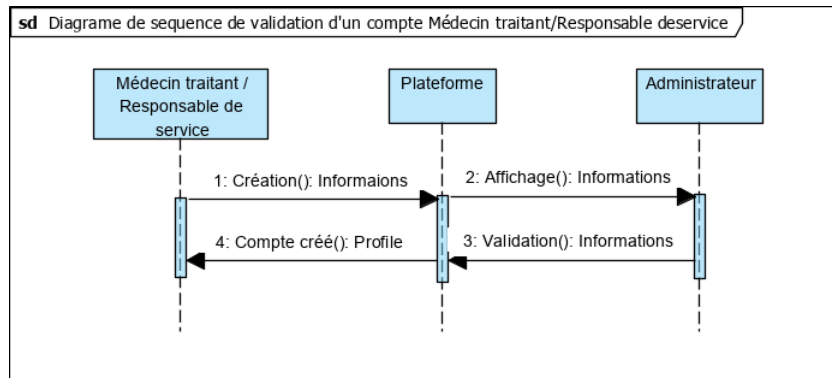


FIGURE 4.1: Diagramme de séquence de compte valide d'un médecin traitant / responsable de service

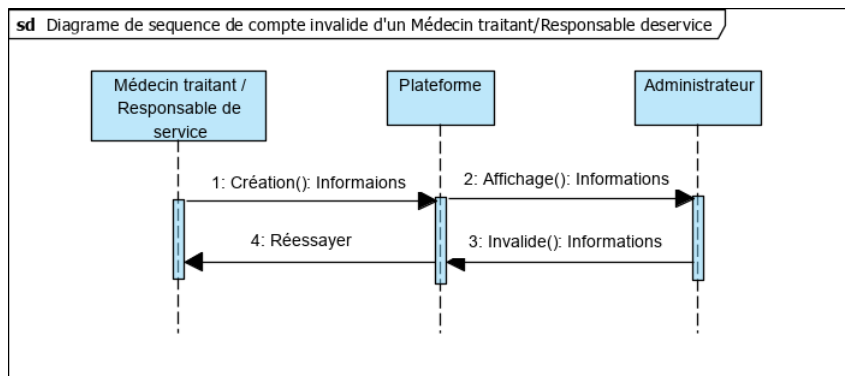


FIGURE 4.2: Diagramme de séquence de compte invalide d'un médecin traitant / responsable de service

#### 4.2.2 Diagramme de séquence de création d'un compte médecin

La figure 4.3 indique le diagramme de séquence de la création d'un compte médecin validé par leur responsable de service, d'autre part la figure 4.4 indique la chronologie des messages d'un compte invalide.

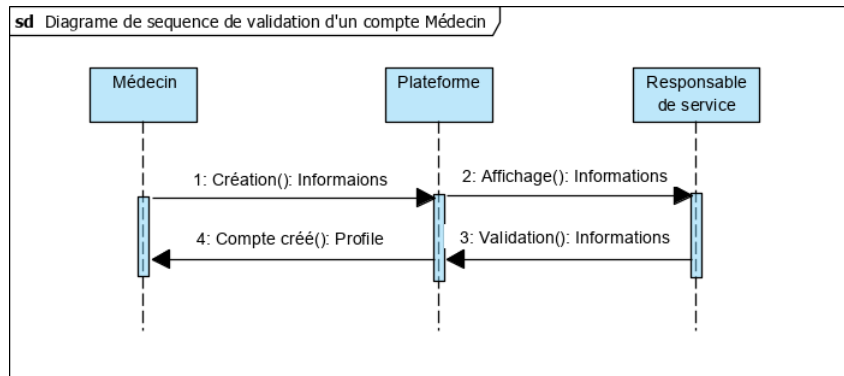


FIGURE 4.3: Diagramme de séquence de compte valide d'un médecin

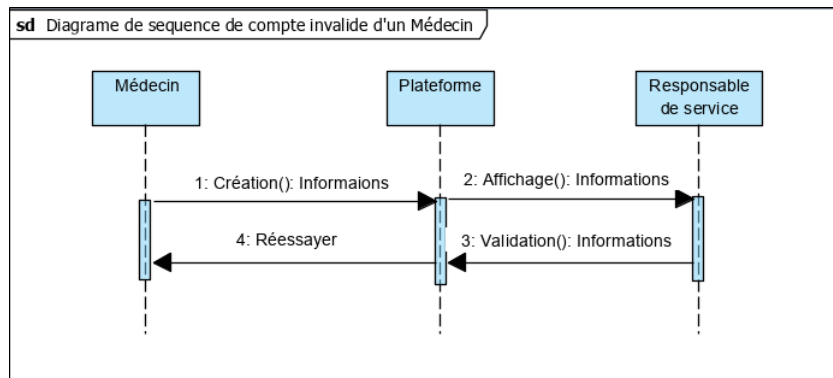


FIGURE 4.4: Diagramme de séquence de compte invalide d'un médecin

### 4.2.3 Diagramme de séquence de création d'un compte patient

La figure 4.5 explique la chronologie des messages entre un médecin traitant choisi par le patient, la plateforme et le nouveau patient en cas de compte valide, et l'autre figure 4.6 en cas de compte invalide.

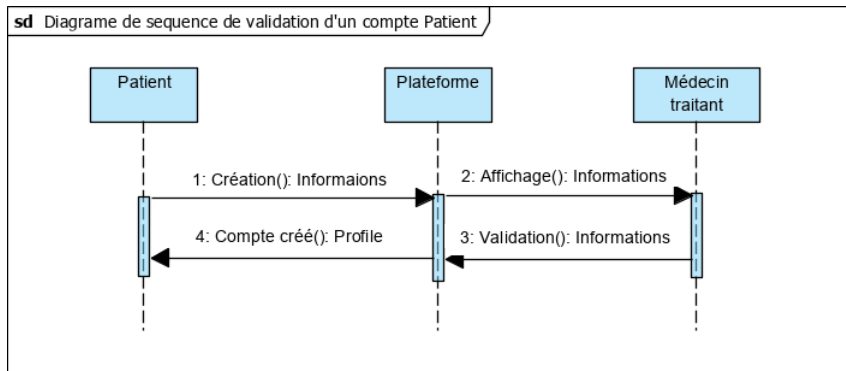


FIGURE 4.5: Diagramme de séquence de compte valide d'un patient

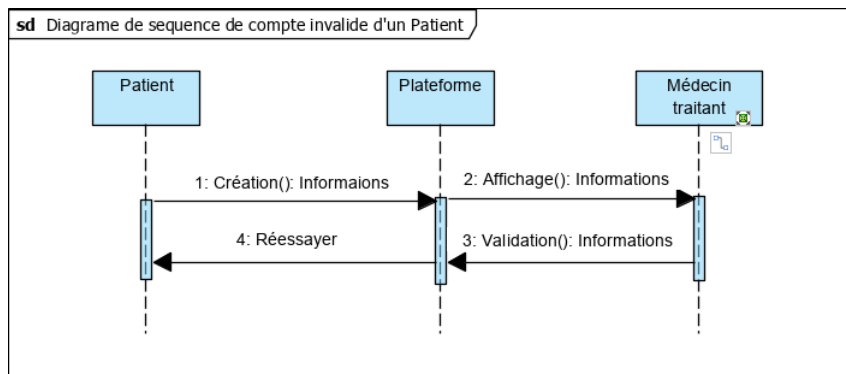


FIGURE 4.6: Diagramme de séquence de compte invalide d'un patient

#### 4.2.4 Diagramme de séquence de prise de rendez-vous

Le scénario présenté à la figure 4.7 explique la chronologie des messages entre le patient, le responsable de service, la plateforme et le médecin, lors d'un rendez-vous valide. Le scénario présenté à la figure 4.8 explique la chronologie des messages entre le patient, le responsable de service et la plateforme, prenant un rendez-vous invalide.

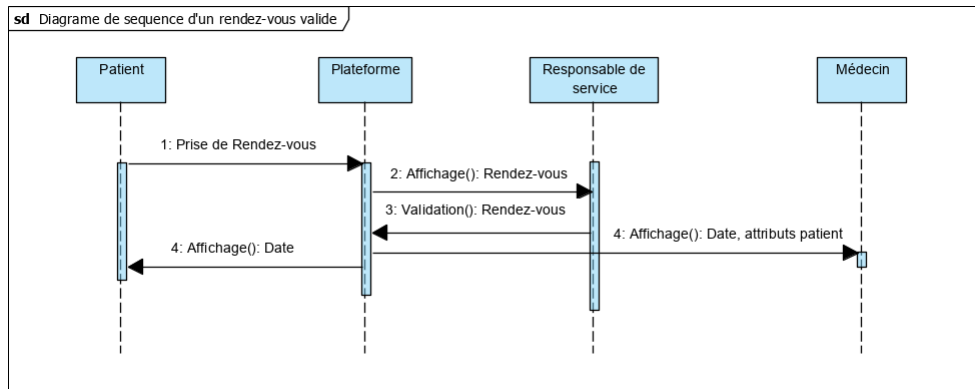


FIGURE 4.7: Diagramme de séquence de prise de rendez-vous valide

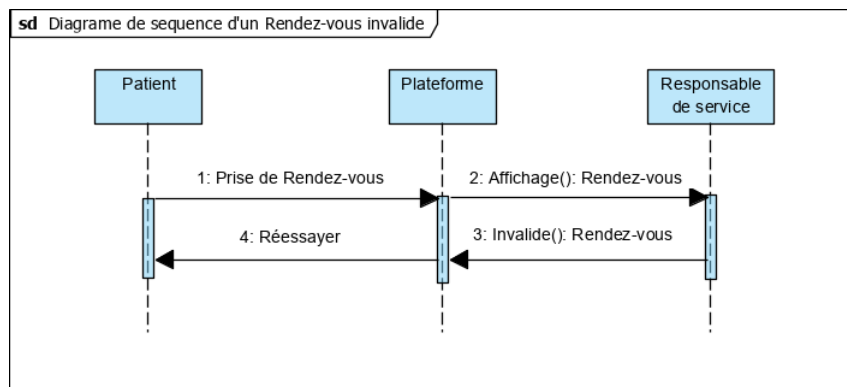


FIGURE 4.8: Diagramme de séquence d'un rendez-vous non valide

#### 4.2.5 Diagramme de séquence des consultations

Le scénario présenté dans la figure 4.9 explique le cas d'une consultation valide, ainsi que la figure 4.10 présente le cas d'une consultation invalide.

Nous prenons en compte les cas de diagnostic invalide selon les exemples suivants :

- Parmi les médicaments qui affectent la chimiothérapie du cancer, nous avons des inhibiteurs calciques, qui sont décrits pour le traitement de l'hypertension artérielle. La figure 4.11 montre la chronologie des messages entre le médecin, la plateforme et le médecin traitant. Le médecin consulte le diagnostic de chimiothérapie anticancéreuse mais le médecin traitant invalide cette consultation car il a déjà d'autres traitements qu'il faut prendre en compte.
- La figure 4.12 montre la chronologie des messages entre le médecin (diagnostiquer un sirop très sucré contenant 60 g de sucre pour 100

ml), la plateforme et le médecin traitant qui invalide le diagnostic pour un diabétique.

- Le médecin, sans le savoir, a prescrit une anesthésie générale mais le patient est thyroïdien, donc la chronologie des messages de la figure 4.13 montre un diagnostic invalide.
- La figure 4.1 montre que le diagnostic invalide par le responsable de service parce que le médecin indique une chirurgie de grossesse sans montrer les causes.

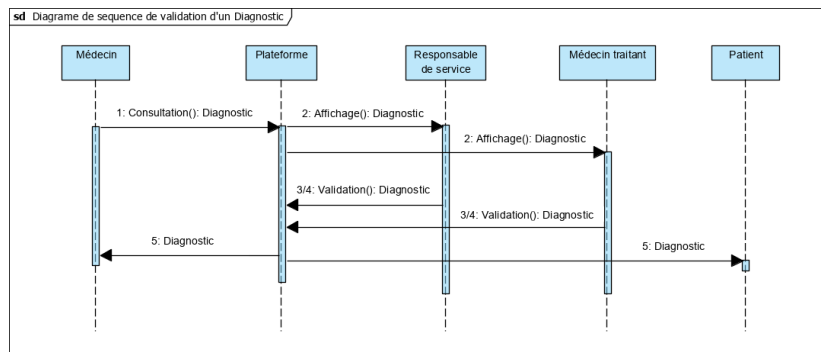


FIGURE 4.9: Diagramme de séquence d’une consultation valide

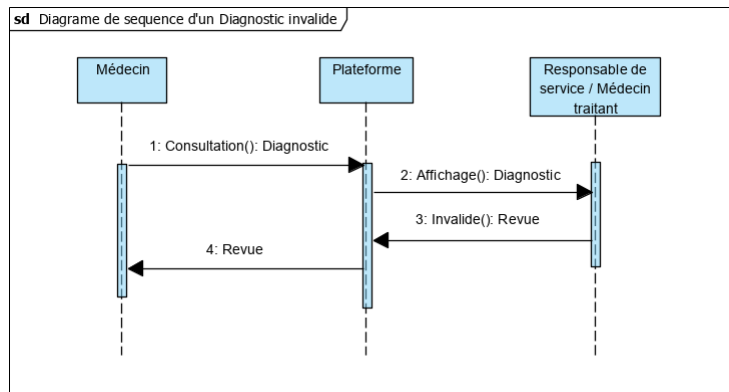


FIGURE 4.10: Diagramme de séquence d’une consultation invalide

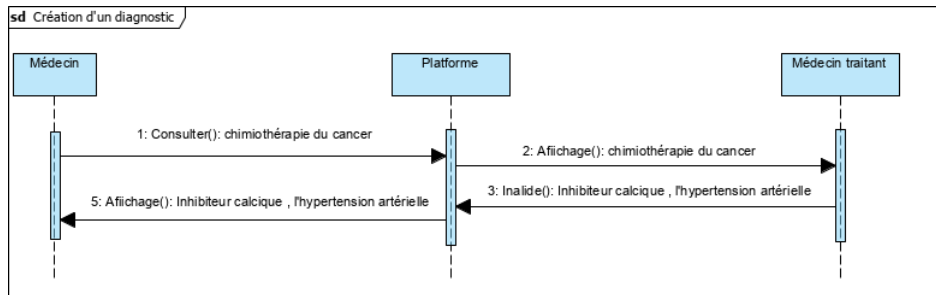


FIGURE 4.11: Exemple 1 d'une consultation invalidée

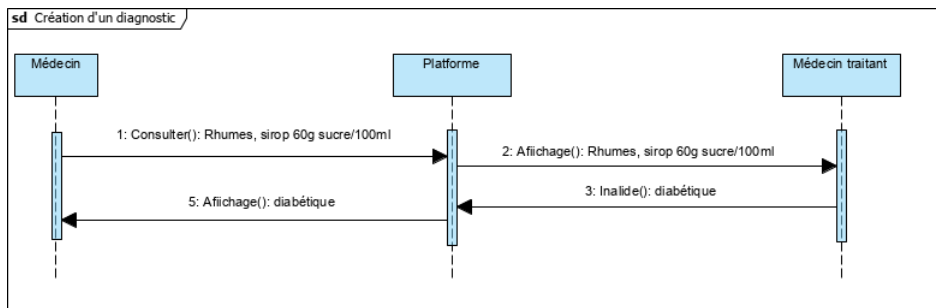


FIGURE 4.12: Exemple 2 d'une consultation invalidée

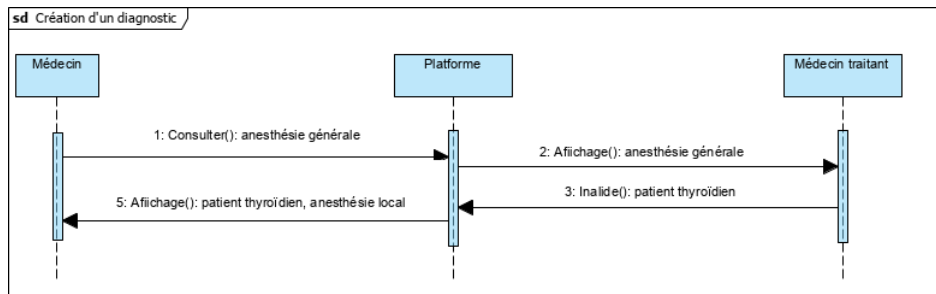


FIGURE 4.13: Exemple 3 d'une consultation invalidée

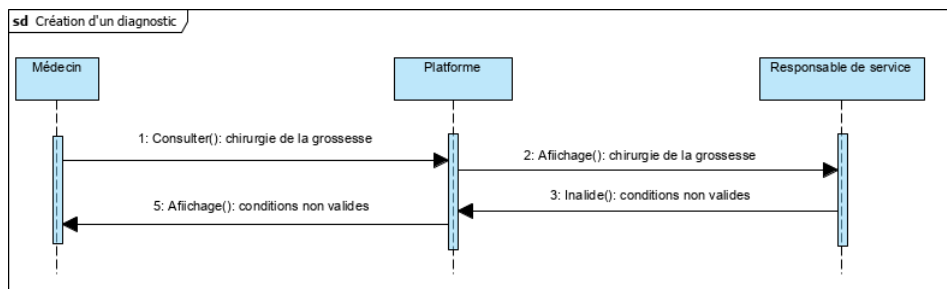


FIGURE 4.14: Exemple 4 d'une consultation invalidée

### 4.3 Outils de développement et langages

Le tableau suivant présente les différents outils utilisés pour l'implémentation de notre système :



#### PyCharm

Assistance intelligente pour Python PyCharm, fournit la saisie automatique de code intelligente, des inspections de code, la mise en évidence d'erreur à la volée et des correctifs rapides, en plus de refactorisations de code automatisées et de riches capacités de navigation[1].



#### Python

Python est un langage de programmation qui vous permet de travailler plus rapidement et d'intégrer vos systèmes plus efficacement. L'index de package Python (PyPI) héberge des milliers de modules tiers pour Python. La bibliothèque standard de Python et les modules apportés par la communauté offrent des possibilités infinies (développement Web et Internet, base de données, accès, interfaces graphiques de bureau, scientifique et numérique, éducation, réseau, programmation, développement de logiciels et de jeux)[2].

## django

### Django

Django est un framework Web Python de haut niveau qui encourage un développement rapide et une conception propre et pragmatique. Conçu par des développeurs expérimentés, il prend en charge une grande partie des tracas du développement Web, vous pouvez donc vous concentrer sur l'écriture de votre application sans avoir à réinventer la roue. C'est gratuit et open source[3].



### Firestore Realtime Database

La base de données en temps réel Firestore est une base de données NoSQL hébergée dans le cloud qui vous permet de stocker et de synchroniser les données entre vos utilisateurs en temps réel[4].



### HTML

HTML est le langage de balisage standard pour les pages Web[5].



### CSS

CSS est un langage qui décrit le style d'un document HTML[5].





## JavaScript

JavaScript est le langage de programmation de HTML et du Web[5].

TABLE 4.1: Outils de développement et langages de programmation

## 4.4 Interfaces de notre plateforme

### 4.4.1 L'administrateur

Un administrateur est une personne qui a une relation directe avec le responsable de service ou le médecin traitant, lorsqu'un nouveau responsable de service crée un compte, ses informations sont affichées sur l'onglet qui s'affiche sur la figure 4.15, la même chose pour un nouveau médecin, comme indiqué sur la figure 4.16.

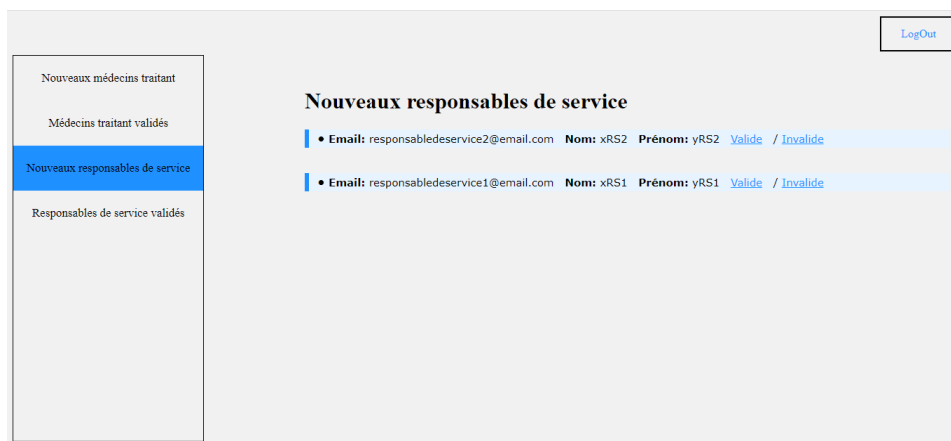


FIGURE 4.15: Nouveaux responsables de services



FIGURE 4.16: Nouveaux médecins traitants

l'administrateur peut valider les nouveaux comptes des responsables de services ou des médecins traitants, tandis que les comptes sont valides, leurs blocs sont créés et leurs informations sont toujours affichées dans l'onglet "Médecins traitants validés" comme indiqué sur la figure 4.18, et l'onglet "Responsables de services validés" comme indiqué sur la figure 4.17 et l'administrateur peut rechercher à un responsable de service ou un médecin traitant avec le code hash.



FIGURE 4.17: Responsables de services validés

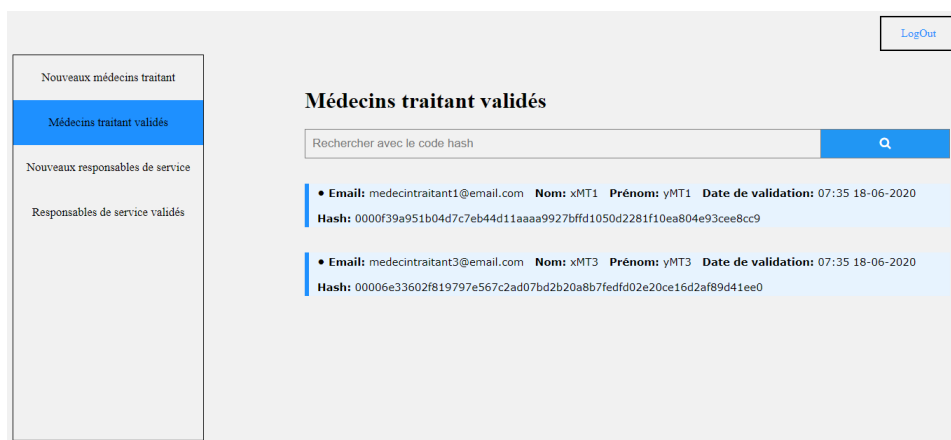


FIGURE 4.18: Médecins traitants validés

#### 4.4.2 Médecins traitants

La figure 4.19 indique l'apparition d'un compte de médecin traitant avant la validation ainsi que la figure 4.20 montre un compte refusé.

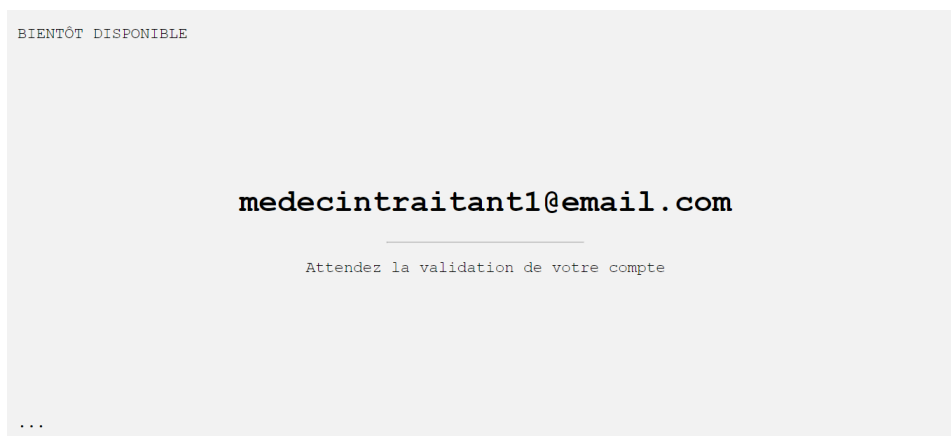


FIGURE 4.19: Exemple d'un compte médecin traitant avant la validation



FIGURE 4.20: Exemple d'un compte médecin traitant refusé

**Note :** Le processus d'attente et d'un compte invalide est le même pour tous les autres acteurs (responsable de service, médecin et patient)

### Profil

Lorsque le compte d'un médecin traitant est validé, le premier onglet qui apparaît après la connexion est son profil comme le montre la figure 4.21.

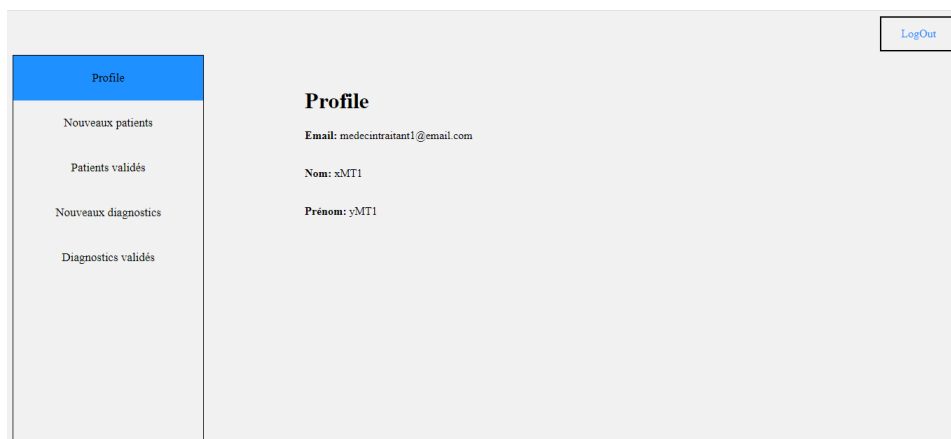


FIGURE 4.21: Profil d'un médecin traitant

### Nouveaux patients

La figure 4.31 montre l'onglet qui affiche tous les nouveaux patients qui choisissent ce médecin traitant lors de leur inscription.



FIGURE 4.22: Nouveaux patients

### Patients validés

Si le médecin traitant approuve le patient, il figurera sur sa liste de patients comme le montre la figure 4.23.



FIGURE 4.23: Patients validés

### Nouveaux diagnostics

Lorsqu'un médecin n'établit le diagnostic sur un patient n, le bloc de diagnostic n'est pas validé uniquement après la vérification du médecin traitant de ce patient et le responsable de service de ce médecin, tandis que les informations de diagnostic apparaissent dans l'onglet illustré sur la figure 4.24.



FIGURE 4.24: Nouveaux diagnostics

Si un médecin traitant ou un responsable de service invalide le diagnostic il est nécessaire d'expliquer les causes de re-diagnostiquer et de faire une revue comme l'illustre la figure 4.25.

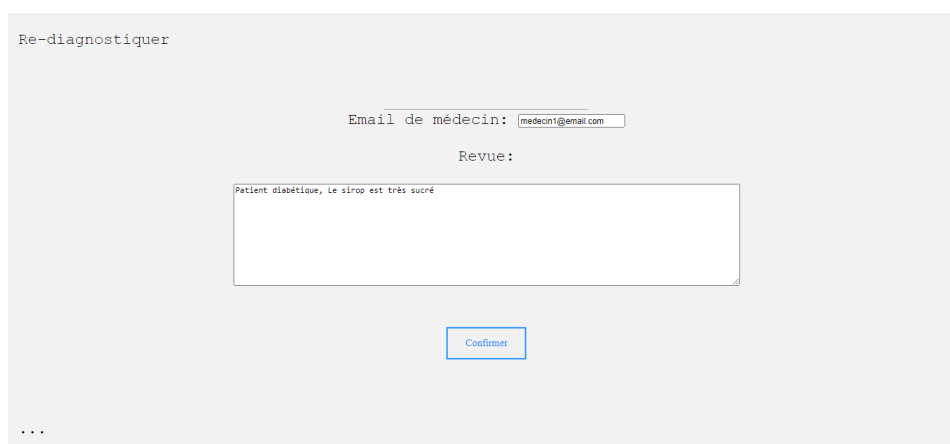


FIGURE 4.25: Exemple d'un revue

## Diagnostics validés

Le dossier du patient doit être examiné avant la validation du diagnostic, après la validation le bloc de diagnostic créé et porte le hachage du médecin et du patient, pour garder la traçabilité de tous les diagnostics d'un patient, et tous les diagnostics posés par un médecin, la figure 4.26 indique l'affichage des codes de hash de diagnostic, patient et médecin après la validation de diagnostic, pour accéder à les informations d'après ces derniers comme indiqué sur les figures 4.27, 4.28, 4.29, ainsi que la recherche peut être effectuée par le hash du patient.

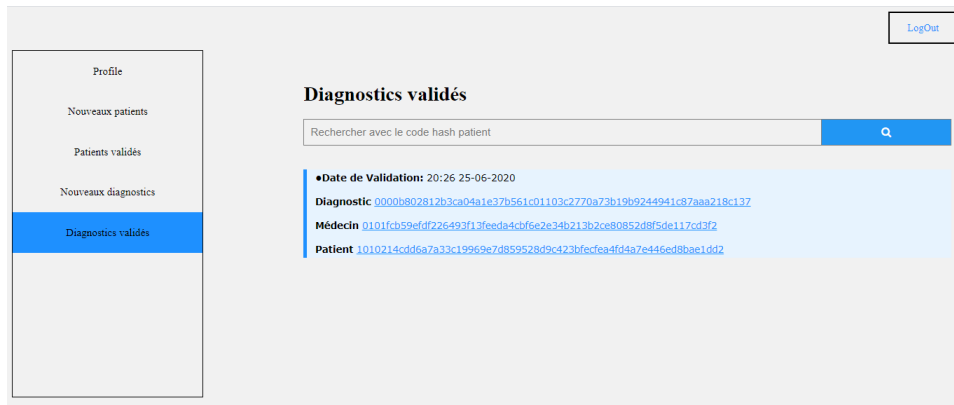


FIGURE 4.26: Diagnostics validés

## Diagnostic

**Type de maladie** Maladie chronique

**Diagnostic** diabétique

**Prix** 1500 DA

FIGURE 4.27: Informations de diagnostic

## Médecin

**Nom** xM1

**Prénom** yM1

**Email** medecin1@email.com

FIGURE 4.28: Informations de patient

# Patient

**Nom** xP1  
**Prénom** yP1  
**Email** patient1@email.com  
**Genre** homme  
**Anniversaire** 1954-01-07

FIGURE 4.29: Informations des médecin

## 4.4.3 Responsables de services

Chaque responsable de services comporte six onglets, à savoir :

### Profil

Comme le montre la figure 4.30.

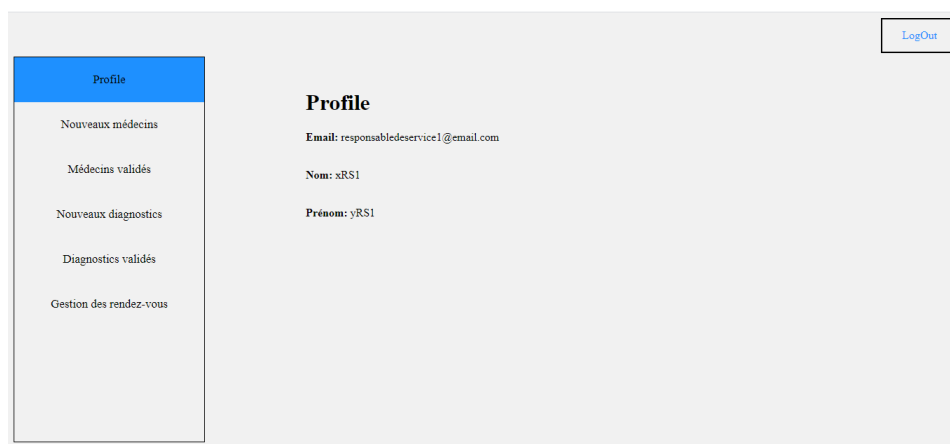


FIGURE 4.30: Profil d'un responsable de service

### Nouveaux médecins

La figure 4.31 montre l'onglet qui affiche tous les nouveaux médecins qui choisissent ce responsable de service lors de leur inscription.





FIGURE 4.31: Nouveaux médecins

## Médecins validés

Si le responsable de service approuve le patient, il figurera sur sa liste de médecins comme le montre la figure 4.32.

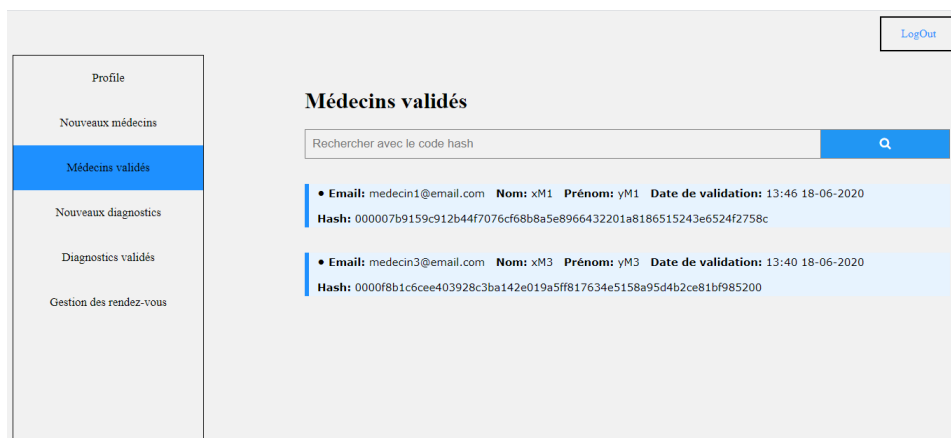


FIGURE 4.32: Médecins validés

## Nouveaux diagnostics

Les informations de diagnostic apparaissent dans l'onglet illustré à la figure 4.33 et validées uniquement après la validation du responsable de service d'un médecin même si le médecin traitant les a validées ou l'inverse.

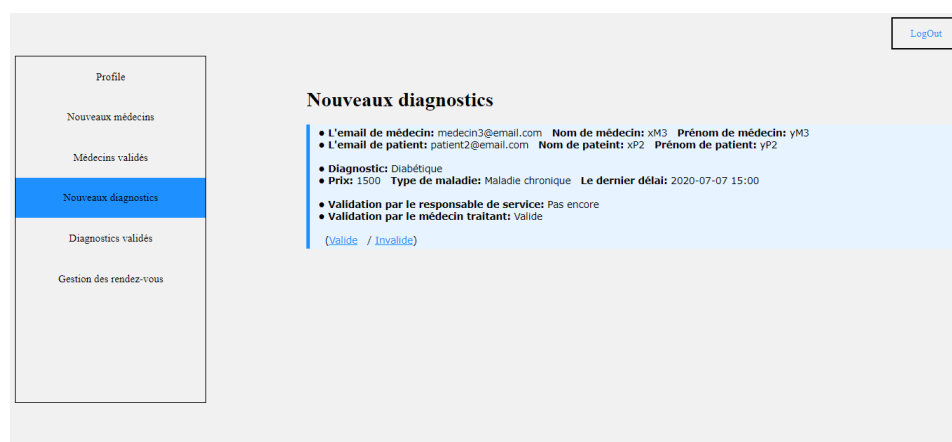


FIGURE 4.33: Nouveaux diagnostics

### Diagnostics validés

La figure 4.34 illustre l'onglet qui affiche tous les code hash qui ont une relation avec les diagnostics effectués par les médecins associés à ce responsable de services, même processus des médecins traitant, tandis que nous pouvons accéder aux informations de diagnostic, médecin, patient. La recherche peut être effectuée par le hash du médecin.

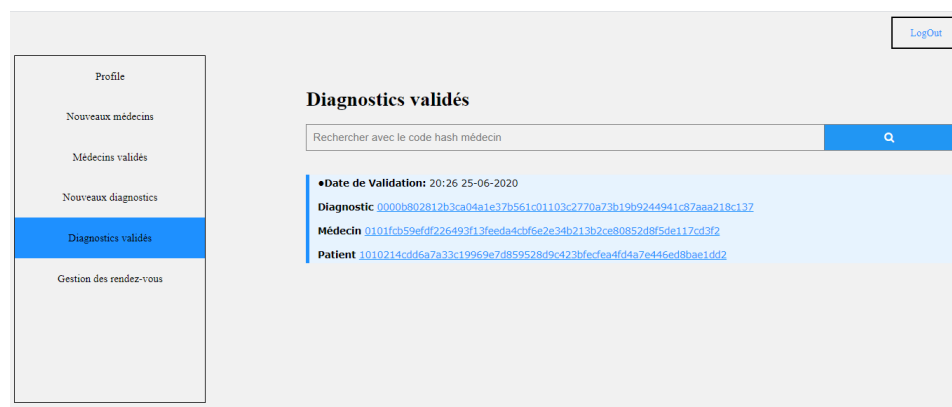


FIGURE 4.34: Diagnostics validés

**La recherche d'un diagnostic** ce processus est très important au côté diagnostique, où chaque responsable peut voir une histoire complète sur leurs médecins, et la même chose pour un médecin traitant avec leurs patients.

Nous l'avons implémenté via le hash pour clarifier davantage l'idée, mais nous pouvons le remplacer par l'email car dans tous les cas, le

programmeur peut accéder au email à partir du hash et ne laisser le hachage que dans la partie programmation.

### Rendez-vous :

cet onglet n'existe que dans le compte du responsable de service associé au médecin choisi par le patient pour valider et donner une date de rendez-vous, comme indiqué dans la figure 4.35.

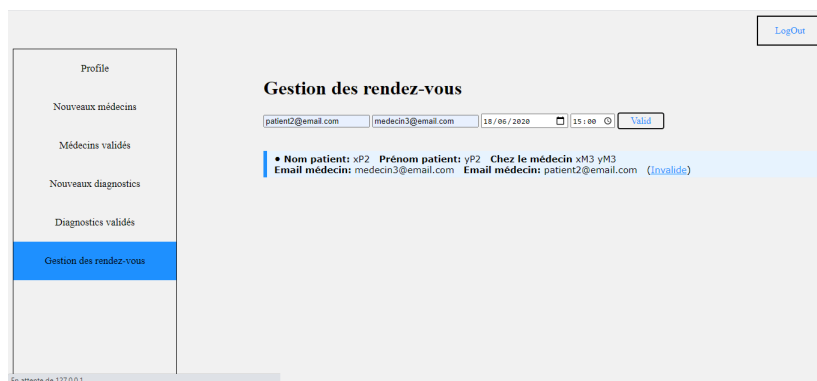


FIGURE 4.35: Gestion des rendez-vous

### 4.4.4 Médecin

Les onglets du compte médecin sont :

#### Profil

La figure 4.36 indique le profil d'un médecin.

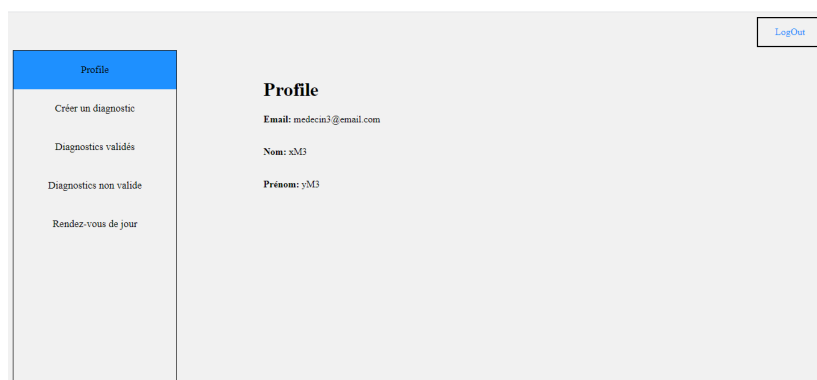


FIGURE 4.36: Profil d'un médecin

## Création de diagnostic

L'onglet montre dans la figure 4.37 est le processus de création d'un diagnostic.

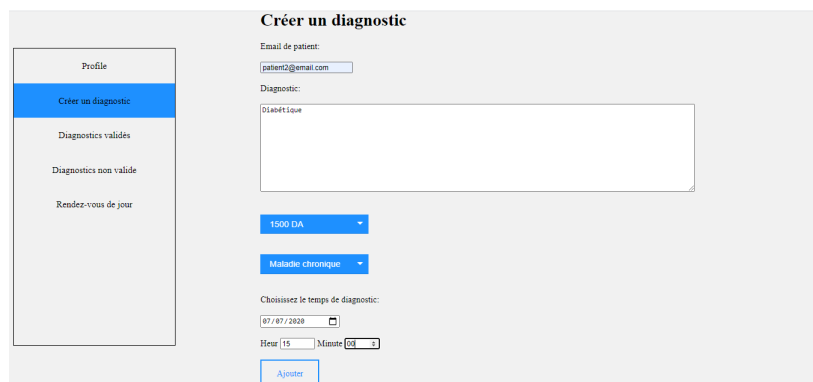


FIGURE 4.37: Création de diagnostic

## Diagnostics validés

Les informations de tous les patients qui consultent un certain médecin apparaissent d'après les code hash diagnostic et patient, comme le montre la figure 4.38, avant la fin de la période de diagnostic qui est un contrat qui expire à la fin de la date spécifiée.



FIGURE 4.38: Diagnostics validés

## Diagnostics invalides

Après le processus d'invalidation par l'un des validateurs (maidecin traitant ou responsable de service) comme indiqué sur la figure 4.25, la revue apparaît pour le médecin dans l'onglet Diagnostics non valide comme indiqué dans la figure 4.39.



FIGURE 4.39: Diagnostics non valide

## Rendez-vous

Après la validation des rendez-vous, ils apparaissent dans cet onglet qui illustre dans la figure 4.40, mais uniquement au date du jour donné.

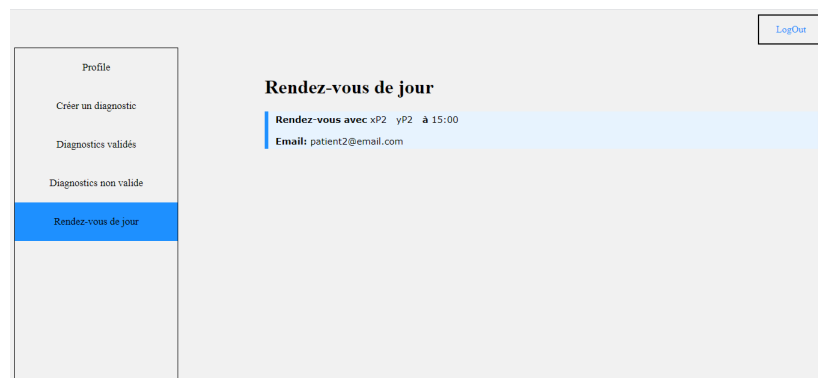


FIGURE 4.40: Rendez-vous par jour

### 4.4.5 Patient

Les onglets du compte patient sont :

#### Profil

La figure 4.41 indique le profil d'un patient.

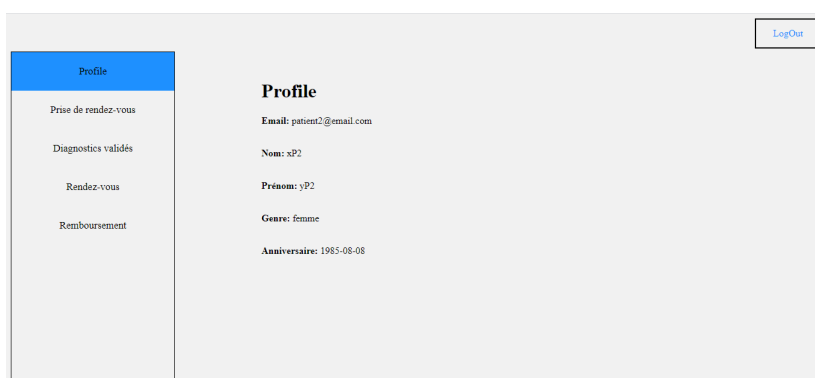


FIGURE 4.41: Profil d'un patient

### Prise de rendez-vous

La figure 4.42 est l'onglet de prise de rendez-vous.

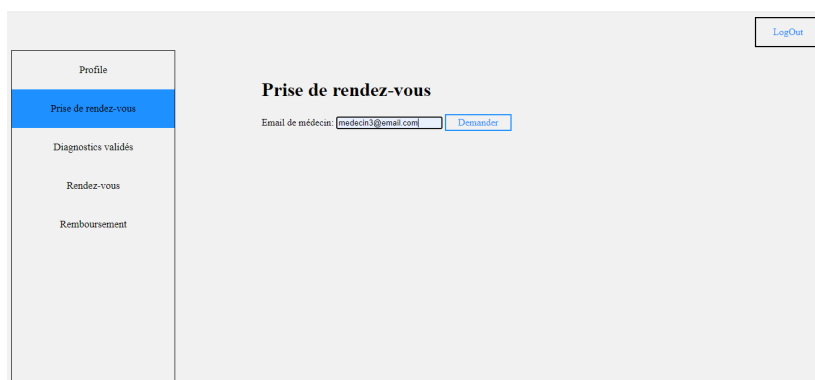


FIGURE 4.42: Prise de rendez-vous

### Diagnostics validés

Les informations de tous les diagnostics posés par un certain médecin sur ce patient apparaissent d'après le code hash diagnostic et médecin comme le montre la figure 4.43, avant la fin de la période de diagnostic qui est un contrat qui expire à la fin de la date spécifiée.



FIGURE 4.43: Diagnostics validés

## Rendez-vous

Après la validation du rendez-vous par le responsable de service, il apparaît dans cet onglet indiqué sur la figure 4.44

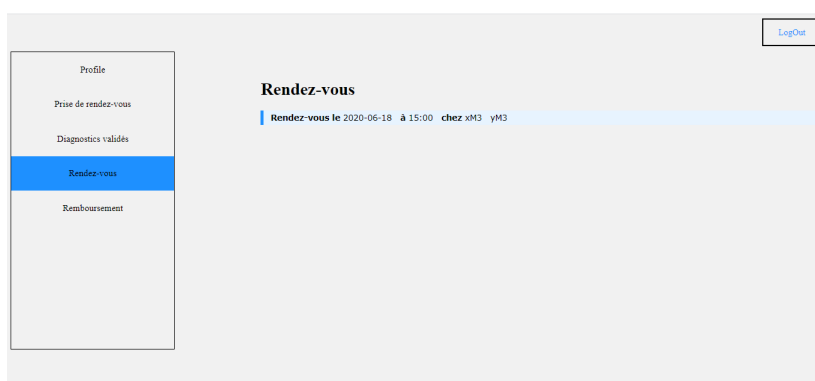


FIGURE 4.44: Rendez-vous

## Remboursement

Le remboursement est une négociation automatique après le type précis de maladie et la validation de diagnostic comme indique dan la figure 4.45

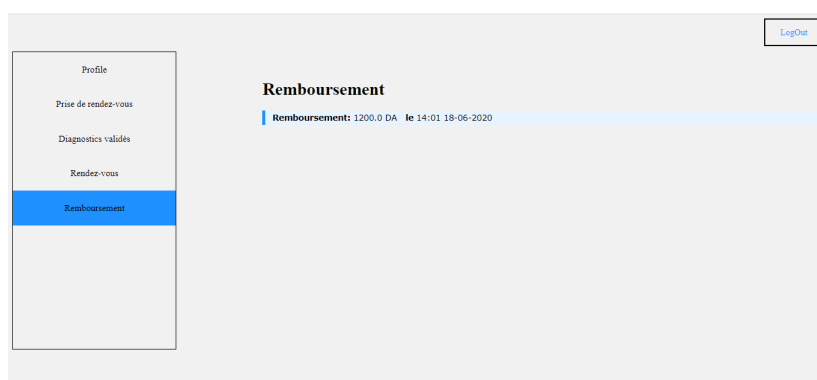


FIGURE 4.45: Remboursement

## 4.5 Stockage off-chain

- Les données sont stockées cryptées à l'aide de la bibliothèque `py-crypto`, il s'agit d'une collection de fonctions de hachage sécurisées (telles que SHA256 et RIPEMD160) et de divers algorithmes de cryptage (AES, DES, RSA, ElGamal, etc.). Le package est structuré pour faciliter l'ajout de nouveaux modules[2]. Nous l'utilisons pour le cryptage RSA et pour la fonction de hachage sha-265, car ce cryptage RSA est variable, chaque fois nous donne une sortie différente pour plus de sécurité.
- Nous fragmentons notre base de données sur différentes bases de données ainsi que sur d'autres données comme indiqué dans le chapitre précédent fragmentation mixte, la figure 4.47 montre un exemple sur la fragmentation des données des blocs médecins et blocs de rendez-vous ainsi qu'un exemple de stockage de données sur d'autres données dans la base de données 5.
- Nous sauvegardons l'intégrité de nos données en dupliquant les données, tandis que le programme copie automatiquement les données brutes au lieu des données modifiées.



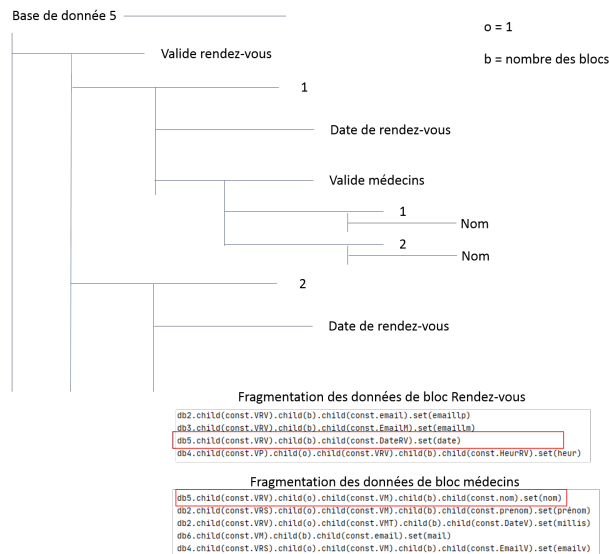


FIGURE 4.46: Fragmetation de données

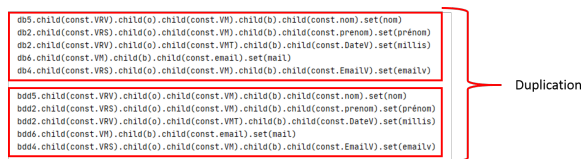


FIGURE 4.47: Exemple de duplication des données

## 4.6 Conclusion

Dans ce chapitre, nous avons présenté les scénarios de notre étude, les outils et langages nécessaires au développement de notre système ainsi que nous avons décrit les différentes interfaces de la plateforme et le stockage.

# Conclusion générale

---

---

Les données médicales sont des données sensibles et vulnérables, d'autant plus que c'est une marchandise qui peut être achetée et vendue, ou cela peut être une provocation pour le patient en violant sa vie personnelle, de sorte que la sécurité est l'élément le plus important dans les systèmes de santé, ainsi que le plus grand défi est de maintenir une sécurité optimale.

Notre système permet de résister ces problèmes en remplaçant le système de sécurité central par un autre qui fournit un stockage dans des bases de données réparties et fragmentées. Afin de renforcer la sécurité des données, nous avons utilisé deux techniques : la première concerne un nouveau modèle de bloc pour augmenter la confidentialité des informations pertinentes et la deuxième concerne l'utilisation de la technique de segmentation des données au moment de l'enregistrement dans le cloud. L'utilisation de blockchains dans les systèmes de santé permet d'assurer plus de transparence et d'efficacité à travers la distribution et l'ordonnancement des données, ainsi d'assurer la pérennité et l'infalsibilité des données.

L'utilisation de contrats intelligents offre une confiance absolue dans l'exécution des accords de notre système par rapport à d'autres systèmes, le caractère transparent, autonome et sécurisé de l'accord élimine toute possibilité de manipulation de partialité ou d'erreur.

Comme perspective, nous espérons intégrer l'IOT et des systèmes d'aide à la décision pour une meilleure gestion des diagnostics, ainsi que utiliser des diagnostics Multimodale avec des techniques de cryptographie mieux adaptées pour les documents multimédias.

# Bibliographie

- [1] <https://www.jetbrains.com/fr-fr/pycharm/>. Consulté juin 2020.
- [2] <https://www.python.org/>. Consulté juin 2020.
- [3] <https://www.djangoproject.com/>. Consulté juin 2020.
- [4] <https://firebase.google.com/>. Consulté juin 2020.
- [5] <https://www.w3schools.com/>. Consulté juin 2020.
- [6] Introduction to information security. [t.ly/eZLZE](https://t.ly/eZLZE). Consulté le 24/04/2020.
- [7] Data privacy, ethics and protection : Guidance note on big data for achievement of the 2030 agenda. [t.ly/2YVGN](https://t.ly/2YVGN), 2017. Consulté le 29 mars 2020.
- [8] Data protection guide (complete). [t.ly/xwJEY](https://t.ly/xwJEY), 2018. Consulté le 29 mars 2020.
- [9] Stages of data vulnerability and the risks. [t.ly/f2Xw](https://t.ly/f2Xw), 2019. Consulté le 16 mai 2020.
- [10] Jamila Alsayed Kassem, Sarwar Sayeed, Hector Marco-Gisbert, Zeeshan Pervez, and Keshav Dahal. Dns-idm : A blockchain identity management system to secure personal data sharing in a network. *Applied Sciences*, 9(15) :2953, 2019.
- [11] Ioannis Boutsis and Vana Kalogeraki. Privacy preservation for participatory sensing data. In *2013 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 103–113. IEEE, 2013.
- [12] M. R. Kabat Chandrasekhar Rao, Amiya Kumar Rath. Cryptography and network security lecture notes. [t.ly/Bk5d7](https://t.ly/Bk5d7). Consulté le 24/04/2020.
- [13] Shawn Dexter. How are blockchain transactions validated? consensus vs validation. <https://is.gd/AN3i5m>, 2018. Consulté le 21/04/2020.
- [14] Ashutosh Dhar Dwivedi, Gautam Srivastava, Shalini Dhar, and Rajani Singh. A decentralized privacy-preserving healthcare blockchain for iot. *Sensors*, 19(2) :326, 2019.

- 
- [15] Rachel L Finn, David Wright, and Michael Friedewald. Seven types of privacy. In *European data protection : coming of age*, pages 3–32. Springer, 2013.
- [16] Kristen N Griggs, Olya Ossipova, Christopher P Kohlios, Alessandro N Baccarini, Emily A Howson, and Thayer Hayajneh. Healthcare blockchain system using smart contracts for secure automated remote patient monitoring. *Journal of medical systems*, 42(7) :130, 2018.
- [17] Sushma Jaiswal. Privacy, protection of personal information and reputation rights : A socio legal analysis. *Our Heritage*, 68(30) :5017–5027, 2020.
- [18] Wolter Martijn Jeroen, Martijn. Privacy and information technology. t.ly/vZRbq, 2019. Consulté le 07 avril 2020.
- [19] Archana Prashanth Joshi, Meng Han, and Yan Wang. A survey on security and privacy issues of blockchain technology. *Mathematical Foundations of Computing*, 1(2) :121–147, 2018.
- [20] Freeha S Khan, Jung Hwan Kim, Robin L Moore, and Lars Mathiassen. Data breach risks and resolutions : A literature synthesis. 2019.
- [21] Shyam Nandan Kumar. Review on network security and cryptography. *International Transaction of Electrical and Computer Engineers System*, 3(1) :1–11, 2015.
- [22] Hongyu Li, Liehuang Zhu, Meng Shen, Feng Gao, Xiaoling Tao, and Sheng Liu. Blockchain-based data preservation system for medical data. *Journal of medical systems*, 42(8) :141, 2018.
- [23] Faiqa Maqsood, Muhammad Ahmed, Muhammad Mumtaz Ali, and Munam Ali Shah. Cryptography : A comparative analysis for modern techniques. *International Journal of Advanced Computer Science and Applications*, 8(6) :442–448, 2017.
- [24] Luci Pangrazio and Neil Selwyn. ‘personal data literacies’ : A critical literacies approach to enhancing understandings of personal digital data. *New Media & Society*, 21(2) :419–437, 2019.
- [25] JEFF PETERS. Data privacy guide : Definitions, explanations and legislation. t.ly/XXewe, 2020. Consulté le 27 mars 2020.
- [26] Shivani Sharma and Yash Gupta. Study on cryptography and techniques. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, 2(1), 2017.
- [27] Nigel P Smart. Hash functions, message authentication codes and key derivation functions. In *Cryptography made simple*, pages 271–294. Springer, 2016.
- [28] Dylan Yaga, Peter Mell, Nik Roby, and Karen Scarfone. Blockchain technology overview. *arXiv preprint arXiv :1906.11078*, 2019.

- [29] Zibin Zheng, Shaoan Xie, Hong-Ning Dai, Weili Chen, Xiangping Chen, Jian Weng, and Muhammad Imran. An overview on smart contracts : Challenges, advances and platforms. *Future Generation Computer Systems*, 105 :475–491, 2020.