

Projet MyPress

Développement d'un système de centralisation d'accès à la presse en ligne

El Abbassi Ilias

Promoteur, Tuteur :

Louis de Viron, Quentin Lurkin

En vue de l'obtention du diplôme de

Master en Sciences de l'Ingénieur industriel

Orientation Informatique

Année académique : 2019 – 2020

Remerciements

Je tiens à remercier tous celles et ceux qui ont participé à la réalisation de mon travail de fin d'études et qui m'ont secondé lors de l'écriture de ce mémoire.

Je voudrais tout d'abord remercier mon tuteur, M. Quentin Lurkin, professeur d'informatique de l'ECAM, pour ses conseils et son suivi

Je remercie également mon promoteur, M. Louis de Viron, fondateur de l'entreprise DataText, pour sa patience et ses conseils, et qui m'a beaucoup appris sur les défis à relever dans le monde du travail

Je tiens à témoigner toute ma reconnaissance à M. Sébastien Combéfis, qui m'a apporté l'aide dont j'avais besoin et de judicieux conseils lors de la réalisation de ce travail de fin d'études, en plus de me suivre régulièrement

Et pour finir, je remercie mes proches pour leur soutien et leurs encouragements.

Abstract

La plupart des sites de presse en ligne suivent un modèle économique identique, basé sur l'abonnement mensuel ou annuel, auquel il faut souscrire afin d'avoir accès aux articles payants proposés par le site de presse.

Cependant, ce modèle pose problème lorsqu'une personne ne souhaite lire qu'un seul article ou quelques articles de manière irrégulière. L'abonnement est alors un modèle plutôt défavorable qui va décourager le lecteur, et qui ne correspond pas toujours à ses habitudes de consommation médiatique.

MyPress est un projet de plateforme qui centralise l'accès à la presse en ligne, en proposant un portefeuille unique à partir duquel un lecteur peut acheter des articles à la pièce auprès des sites de presses partenaires.

Ce travail de fin d'études concerne la mise en place de la plateforme informatique supportant le projet, depuis l'architecture jusqu'à l'implémentation d'un prototype fonctionnel. Celui-ci repose sur une architecture Web MVC et un flux d'informations entre les différents acteurs du projet.

Une base de données relationnelle permet de stocker les informations nécessaires et de réaliser des requêtes statistiques. La sécurité du projet et de son API¹ est réalisée par plusieurs concepts tels que le Json Web Token, le hashage sécurisé des mots de passe ou une protection face aux injections SQL.

Le premier partenaire de MyPress est Alter Echos, un journal indépendant qui analyse des problématiques sociales. Ce dernier permet de tester les fonctionnalités de MyPress et d'établir le modèle économique du projet.

¹ Application Programming Interface

CAHIER DES CHARGES RELATIF au TRAVAIL DE FIN D'ETUDES de

El Abbassi Ilias inscrit(e) en Informatique

Année académique : 2019-2020

Titre provisoire :

Développement d'un projet web de service de connexion

Objectifs à atteindre :

- Développement d'une application web commencée durant la période de stage
- Mise en place du projet sur la plateforme des potentiels clients

Principales étapes :

- Amélioration du site web
- Sécurisation du projet
- Travailler sur la compatibilité du projet avec les plateformes des clients
- Mise en place du projet
- Amélioration continue
- Etude de marché sur la faisabilité du projet

Fait en trois exemplaires à Bruxelles, le 18/11/2019

L'Etudiant

Le Tuteur

Le Promoteur

Nom-prénom :

EL ABBASSI
ILIAS

Nom-prénom :

Nom-prénom :

de Virel Louis

Département/Unité

Société

DATAEXT...SRL

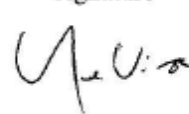
Signature



Signature



Signature



Coronavirus

D'abord, le coronavirus et la mise en place du confinement ont eu un impact sur mon travail de fin d'études ainsi que sur sa rédaction, principalement sur l'organisation.

Ensuite, il a limité les échanges avec AlterEchos, le partenaire du projet MyPress. Il était normalement prévu que nous allions discuter avec les membres d'AlterEchos afin de se mettre d'accord sur l'interconnexion du projet et que je travaille avec leur développeur pour intégrer les scripts qui vont permettre une coopération entre leur site de presse et notre plateforme.

La collaboration avec ce partenaire est pour le moment en attente.

Table des matières

1. Introduction	8
2. Contexte	10
2.1. Problématique.....	10
2.2. Étude de marché	11
3. Entreprise	15
4. Environnement de développement	17
5. Modélisation du problème	18
5.1. Vue d'ensemble du système	18
5.2. Base de données.	19
5.2.1. Données et systèmes	19
5.2.2. Modèle entité-association	21
5.3. Force de l'architecture	22
6. Design du prototype	23
6.1. But	23
6.2. Schéma relationnel du prototype	24
6.3. Architecture MVC.....	26
6.4. Analyse fonctionnelle.....	27
6.4.1. Identification des fonctionnalités.....	27
6.4.2. Gestion des connexions et des articles	28
6.4.3. Token de sécurité API.....	31
6.5. Transfert des données et Cookie	32
6.5.1. Connexion du média	32
6.5.2. Connexion utilisateur	33
6.5.3. Vérification des accès aux articles.....	34
6.5.4. Achat d'un article	35
6.6. Authentification	36
6.7. Rechargement de crédits	37
7. Test.....	38
7.1. Stratégie	38

7.2.	Démarche de test	38
7.3.	Site fictif : BelgiumPost	40
7.4.	Site externe	41
8.	Sécurité.....	42
8.1.	Attaque et danger	43
8.2.	Analyse de risque	44
8.2.1.	Analyse de risque : Base de données MyPress	46
8.2.2.	Analyse de risque : API MyPress	47
8.3.	Certificat HTTPS.....	48
8.4.	Attaque XSS	48
8.5.	Injection SQL	49
8.6.	Hashage des mots de passe	50
8.7.	JSON Web Token	52
9.	Analytics	54
10.	Perspective	55
11.	Conclusion.....	56
12.	Bibliographie	58
13.	Annexes.....	60
13.1.	Accueil et ses boutons.....	60
13.2.	Inscription.....	62
13.3.	Lecteur connecté.....	62

1. Introduction

Avec l'avènement de l'informatique et la démocratisation des ordinateurs pour la population, nombreux sont les groupes de presse et rédacteurs d'articles indépendants à avoir profité de cette avancée technologique pour partager leurs travaux sur différents sites de presse, blog ou tout autre type de média.

L'accès à l'information s'est grandement répandu, et tout le monde ou presque, que ce soit adolescents, jeunes adultes ou personnes plus âgées s'intéresse à la presse et à ce qui se passe à travers le monde.

De plus, les consommateurs de presse ne sont plus limités à lire que des articles provenant de leur ville ou de leur pays. Ils ont à présent accès à de l'information provenant de médias du monde entier.

Dans les premières années d'Internet, les médias ont commencé à diffuser gratuitement leurs articles sur leurs premiers sites web. Puis, se rendant compte de la perte de revenu considérable que cela engendrait, ils ont ajouté de la publicité sur ces articles. Ensuite, toujours à la recherche d'un modèle économique pertinent, ils en sont revenus à une formule traditionnelle, celle de l'abonnement, passant de l'abonnement papier à une version numérique. Or, cette formule ne convient pas à tous les publics. En effet, les lecteurs appartenant à la génération Y² préfèrent consommer différents médias, plutôt que de s'attacher à une seule marque, comme le faisaient leurs prédécesseurs. Or, aucune offre ne permet aujourd'hui de consommer des articles de façon panachée.

Constatant ce vide, l'entreprise DataText, spécialisée en data science et text mining, souhaitant diversifier ses activités, est en train de mettre en place le projet MyPress. Le but de ce projet est de centraliser l'accès à la presse et de proposer une offre alternative, consistant en un compte unique qui permet d'acheter des articles à la pièce chez différents sites de presse partenaires du projet.

C'est dans le cadre de ce projet qu'est réalisé mon travail de fin d'études qui a pour objectif de réaliser un prototype fonctionnel qui respecte les demandes de l'entreprise, afin de démontrer la faisabilité du projet.

² Personnes nées entre le début des années 1980 et la fin des années 90

Il faut alors imaginer une architecture du système adaptée aux attentes de DataText. Celle-ci doit être fonctionnelle pour tester différents types d'utilisations et doit être sécurisée afin de protéger les lecteurs, mais aussi les différents médias. Elle doit être flexible et facilement modifiable pour évoluer avec l'ajout de nouvelles fonctionnalités ou la prise en compte de nouvelles contraintes, et claire et documentée, afin de permettre à l'entreprise de finaliser le projet en utilisant un autre langage de programmation ou un autre type d'implémentation.

Dans ce travail, nous commençons par expliquer la problématique associée à la presse en ligne, ainsi que des défis de la création du projet MyPress.

Ensuite, nous expliquons la modélisation du problème avec tous les composants majeurs qui seront nécessaires à la réalisation du projet, suivi du design du prototype et des choix technologiques adoptés. Un chapitre présente ensuite les différents moyens mis en place afin de tester les fonctionnalités du prototype. Après, un autre expose les vulnérabilités du système et la sécurisation de celui-ci. Puis, une section « analytics » met en évidence différentes requêtes et statistiques utiles. Enfin, un chapitre présente les perspectives du projet et la conclusion de ce travail de fin d'études.

2. Contexte

La création du projet MyPress répond à une problématique propre à la génération Y : leur mode de consommation des médias ne correspond plus à l'offre proposée par la plupart des organes de presse aujourd'hui : la formule d'abonnement.

2.1. Problématique

Les sites de presse en ligne proposent dans leur grande majorité un accès à leurs articles suivant le modèle économique de l'abonnement. Ce modèle économique est avantageux pour le média puisqu'il fidélise le consommateur, et l'incite à lire d'autres articles, puisque l'abonnement lui donne accès à tous les articles qu'il propose.

Ce type de formule dissuade également le client de lire des articles sur d'autres sites de presse, puisqu'il devra payer un autre abonnement s'il veut consulter les articles d'autres médias. Cependant, il pose problème pour de nombreux potentiels clients, et ce pour deux principales raisons.

D'une part, beaucoup de personnes n'ont pas les moyens ou ne souhaitent pas payer un abonnement à un journal. D'après un article de « Presse-citron », les internautes ne sont plus intéressés par un abonnement à chaque fournisseur d'articles de presse, et sont plutôt intéressés par un système économique « à la demande » [14].

D'autre part, les nouvelles générations préfèrent pouvoir lire des articles sur plusieurs médias, plutôt que de se fixer sur un seul organe de presse. Ils ne sont intéressés que par quelques articles par média, et sont donc peu intéressés par un système d'abonnement. En effet, comme décrit dans un article de la « RTBF », selon une étude de Médiamétrie, les jeunes ont tendance à s'informer à travers les réseaux sociaux qui proposent des articles variés, au lieu d'accéder directement à un site de presse. Ainsi, près de 70% des jeunes s'informent d'abord en passant par les réseaux sociaux. [15]

C'est dans ce contexte que se situe le projet MyPress : l'objectif est de donner un accès payant à la presse en ligne, avec un système de portefeuille virtuel qui permettra au consommateur d'acheter à la pièce des articles de différents fournisseurs.

Cette plateforme fonctionnera de manière transparente pour l'utilisateur et prendra en charge la gestion des flux financiers, que ce soit les flux du lecteur vers la plateforme, ou de la plateforme vers les fournisseurs d'articles de presse.

Du point de vue de l'utilisateur, un bouton sera intégré sur le site de presse et se chargera de rendre compatible le site avec MyPress.

2.2. Étude de marché

Afin d'évaluer l'intérêt que pourraient porter les consommateurs à cette nouvelle façon d'accéder à la presse en ligne, une petite étude de marché a été mise en place avec un formulaire Google Form. Cette étude a été partagée sur les réseaux sociaux et a récolté un total de 84 répondants.

Plus de 80% des répondants ont entre 20 et 35 ans. Sachant que cette génération est le cœur de cible de MyPresse, cette surreprésentation n'est pas un problème pour nous.

Quelle est votre situation professionnelle ?

84 réponses

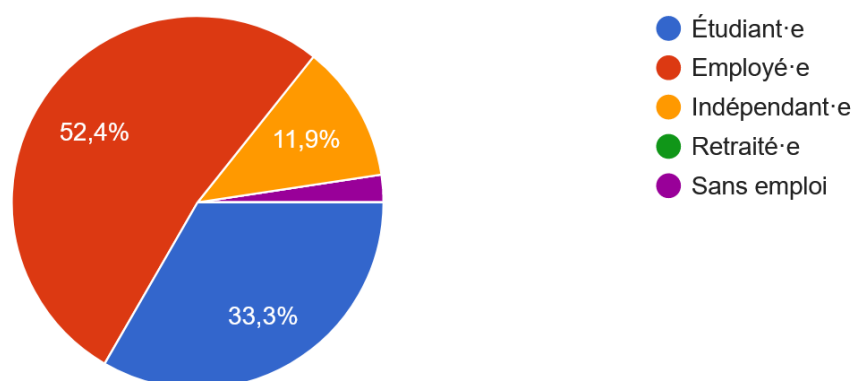


Figure 1 : Étude de marché – Situation professionnelle

La majorité des répondants sont employés, suivi par les étudiants et ensuite les indépendants

Êtes-vous abonné-e à un (ou plusieurs) journaux "papier" ?

84 réponses

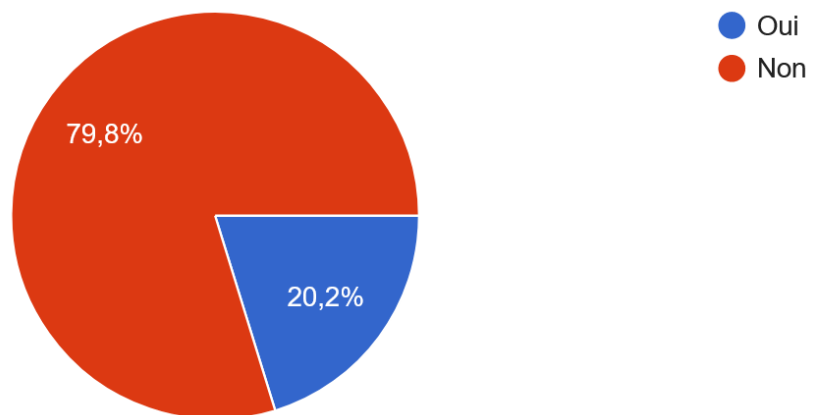


Figure 2 : Étude de marché – Abonnement classique

Près de 80% des répondants ne sont pas abonnés à un média de presse classique, la presse au format papier n'a donc pas beaucoup de succès parmi nos répondants

Possédez-vous un abonnement "digital only" à un ou plusieurs journaux en ligne ?

84 réponses

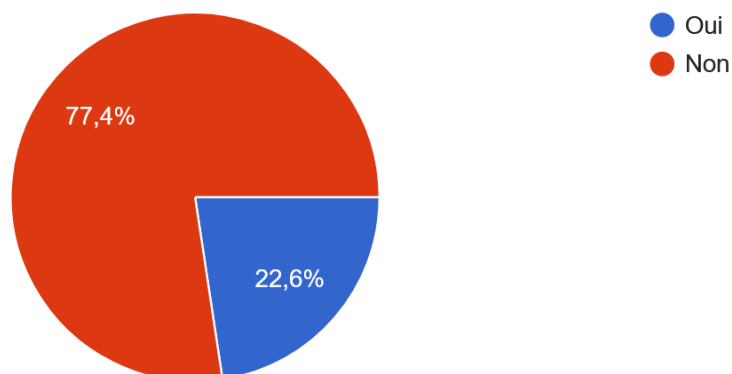


Figure 3 : Étude de marché – Abonnement en ligne

Plus de 75% des répondants ne sont pas abonnés à un journal en ligne

Ces deux derniers graphiques tendent à montrer que les répondants ne semblent pas intéressés par un abonnement à un média de presse

Combien d'articles en ligne (gratuits ou payants) consultez-vous par jour ?

84 réponses

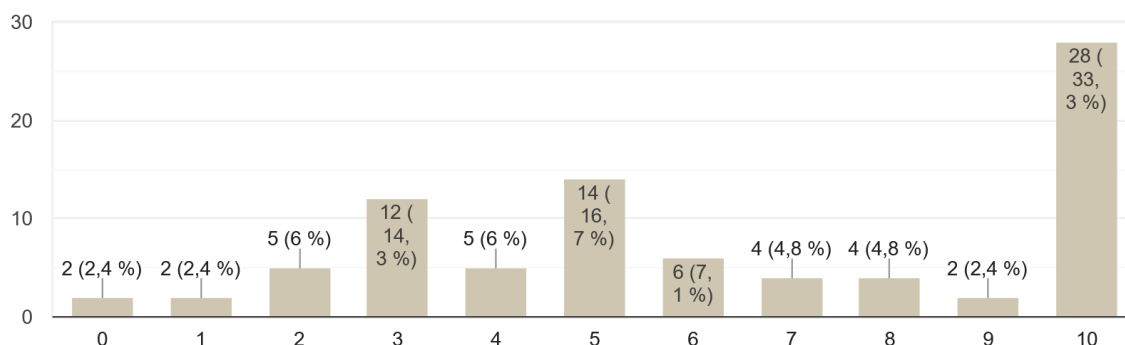


Figure 4 : Étude de marché – Nombre d'articles lus par jour

On voit par contre que les répondants sont intéressés par la presse en ligne, gratuite ou payante, qu'ils consultent régulièrement : près de 70% consultent au moins 5 articles par jour

Êtes-vous intéressé-e par l'achat d'articles à la pièce ?

84 réponses

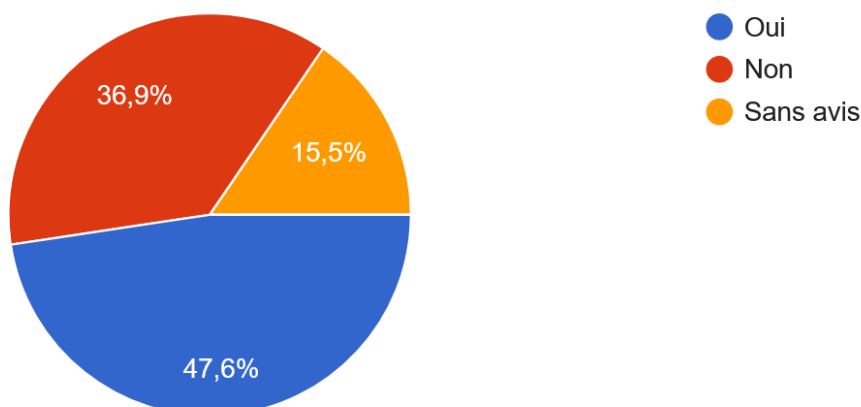


Figure 5 : Étude de marché – Achat à la pièce

Près de la moitié des répondants sont intéressés par un autre modèle économique que l'abonnement, permettant des achats d'articles à la pièce. Cela tend à justifier la pertinence de l'outil sur lequel nous travaillons.

On constate donc qu'une part jeune de la population est intéressée par la presse en ligne qu'ils consultent régulièrement, bien qu'ils soient peu nombreux à posséder un abonnement à un média de presse. De plus, une partie de ceux-ci sont intéressés par un modèle d'achat à la pièce, au lieu de l'abonnement. Cependant, l'étude de marché n'est pas vraiment représentative dû aux nombres restreints de répondants.

DataText a décidé de développer le projet MyPress, afin de faciliter l'accès à la presse en ligne avec un autre modèle économique, visant prioritairement les personnes intéressées par la presse en ligne, mais qui n'ont jamais fait le pas de contracter un abonnement numérique.

Le projet prend la forme d'un site internet, côté client, sur lequel ils s'inscrivent et peuvent acheter des crédits afin de débloquent des articles sur les sites de presse partenaires du projet. Les sites de presse, quant à eux, devront intégrer des parties de code au sein de leur infrastructure afin de permettre la coopération entre leur site et un serveur d'API, et la partie backend du projet MyPress.

Le projet présente deux grands défis :

- Concevoir le projet MyPress en partant de zéro, en passant par l'architecture et la base de données, jusqu'à la sécurisation de la plateforme
- Développer un script prêt à l'emploi et facile à importer pour les différents sites de presse

3. Entreprise

DataText est une entreprise belge fondée par Louis de Viron en 2019. Celle-ci est basée à Bruxelles.

Voici les informations légales concernant l'entreprise Datatext :

- Numéro d'entreprise : 0721.433.837
- Date de création : 21 février 2019
- Dénomination : DataText SPRL
- Adresse du siège : Rue Emile Wittmann 36, Boîte ET02
1030 Schaerbeek
- Situation juridique : Situation normale
- Forme légale : Société privée à responsabilité limitée
- Code d'activité :
 1. 62010 - Programmation informatique
 2. 62020 - Conseil informatique
 3. 62090 - Autres activités informatiques

DataText est une société de conseils, avec une expertise en Data Science et en Natural Language Processing, dont le but est d'aider d'autres entreprises à traiter et interpréter les données internes ou externes grâce à des techniques avancées d'intelligence artificielle, avec un accent sur les données non structurées.

L'entreprise aide principalement de jeunes entreprises qui souhaitent prendre une orientation «data» telle que Koopol ou Novable. Elle travaille également pour des organismes de presse, en proposant des outils d'analyse et de visualisation de données textuelles à caractère politique (RTBF, Wilfried) et participe aussi à des projets data pour des organismes publics tels que la Commission européenne.

DataText cherche également à proposer des produits innovants dans le secteur de l'informatique, en réfléchissant aux besoins des entreprises et la faisabilité de plusieurs projets. C'est dans ce cadre que le projet MyPress a vu le jour.

Pour le développement de ce prototype, j'ai principalement travaillé seul et mon travail s'est structuré en deux parties. D'une part, j'ai étudié la faisabilité du projet. D'autre part, j'ai créé une plateforme fonctionnelle afin de proposer un prototype pour le projet. L'entreprise s'est chargée de mettre en place le projet (rencontre de partenaires, etc.) et M. Combéfis, partie prenante du projet MyPress, m'apportait un soutien sur les aspects techniques.

4. Environnement de développement

Le projet est organisé selon une approche de méthodologie agile. Une réunion permet de définir un bloc de fonctionnalités qui constitue l'objectif à court terme, qui doit être implémenté avant la réunion suivante. Durant cette dernière, le prototype, muni des nouvelles fonctions, est présenté à travers une démonstration, et un autre bloc de fonctionnalités est défini.

Le prototype réalisé dans le cadre de ce travail de fin d'études est programmé en PHP 7.2.18, sans utilisation de framework. Sa base de données est gérée avec MySQL.

Il est conçu grâce à « WampServer », qui est une plateforme de développement Web capable de lancer un programme PHP localement.

Le projet est régulièrement envoyé dans le service de gestion de développement logiciel « BitBucket » qui fonctionne avec un système de gestion de versions Git, ce qui permet de voir son avancée et de revenir à un état antérieur si une étape de production pose problème.

5. Modélisation du problème

Afin de déterminer quels seront les éléments qui devront être utilisés afin de résoudre notre problème, il nous faut d'abord le modéliser. Cette modélisation nous permet de comprendre quelles seront les données qui seront stockées et quelle infrastructure il faudra mettre en place afin de pouvoir réaliser et déployer une application adaptée au problème.

5.1. Vue d'ensemble du système

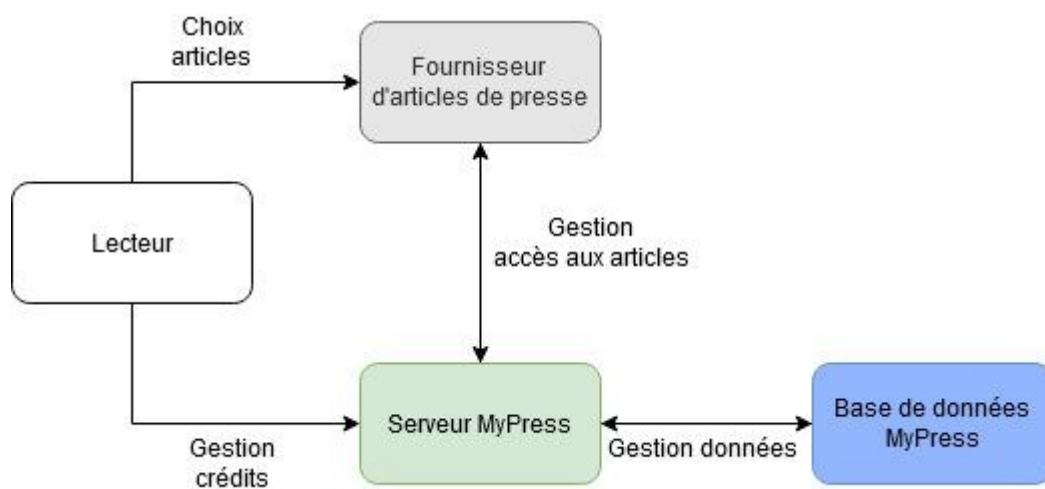


Figure 6 : Vue d'ensemble de l'architecture

Les composants suivants permettent de constituer le prototype :

- La base de données qui stockera les informations clients, achats, etc.
- Le serveur du projet qui va dialoguer avec la base de données afin de vérifier quel client a accès à quel article, combien de crédits possèdent le client, etc.
- Le nom de domaine permettant aux lecteurs et médias d'accéder à la plateforme du projet et de gérer les connexions en Back-end.
- Un front-end pour l'affichage des fonctionnalités aux consommateurs et médias

D'autres composants dépendent de l'implémentation et peuvent permettre de simplifier l'architecture du projet, ou d'augmenter la sécurité de celui-ci. C'est le cas de l'API qui va servir à dialoguer entre les sites de presse et le projet.

Le lecteur choisit un article qu'il souhaite lire sur les fournisseurs d'articles de presse en ligne. Afin de savoir si celui-ci a le droit de lire l'article, le fournisseur envoie une requête au serveur MyPress, qui questionnera sa base de données afin de savoir si le lecteur a déjà acheté cet article. Si oui, l'accès est accordé, sinon l'accès est refusé.

En cas de refus, le lecteur peut utiliser les crédits de son compte MyPress afin d'acheter l'accès à l'article qu'il souhaite lire. S'il n'a pas assez de crédits, il recharge ceux-ci directement sur la plateforme de MyPress

5.2. Base de données.

Les informations doivent être stockées afin de gérer la plateforme, les connexions à celle-ci et gérer les accès de lecteurs aux articles des médias partenaires

Elles seront stockées dans une base de données dont il faut choisir le système de gestion.

5.2.1. Données et systèmes

Un système de gestion de base de données doit être mis en place principalement pour gérer les comptes des utilisateurs et des partenaires, et pour stocker les transactions.

Certaines informations devront être gardées, quel que soit la façon dont le projet sera implémentée :

- Données permettant d'identifier un lecteur
- Données permettant d'identifier un fournisseur d'article de presse
- Informations de connexion pour lecteur et fournisseur
- Données d'identification d'un article
- Crédits

D'autres informations dépendront de l'intégration avec le média partenaire :

- Prix de l'article
- Gestion des flux financiers entre MyPress et le site de presse

D'autres encore seront enregistrées afin de faciliter l'utilisation de la base de données et permettre des requêtes statistiques, telles que :

- Date d'achat d'un article
- Date de lecture d'un article
- Transactions d'achat de crédits

Il existe différents systèmes de gestion de données (Relationnel, orienté-document, orienté-graphe, etc.). La plupart sont utilisables pour ce type de projet, mais le plus simple serait de choisir une base de données de type relationnel (SQL) étant donné que c'est un système standard, très documenté, et qui respecte la propriété d'atomicité (une transaction incomplète est annulée, et l'état de la base de données est remis à l'état initial). De plus, une base de données NoSQL est intéressante lorsqu'il s'agit de traiter un très grand volume de données, ce qui n'est pas le cas ici.

5.2.2. Modèle entité-association

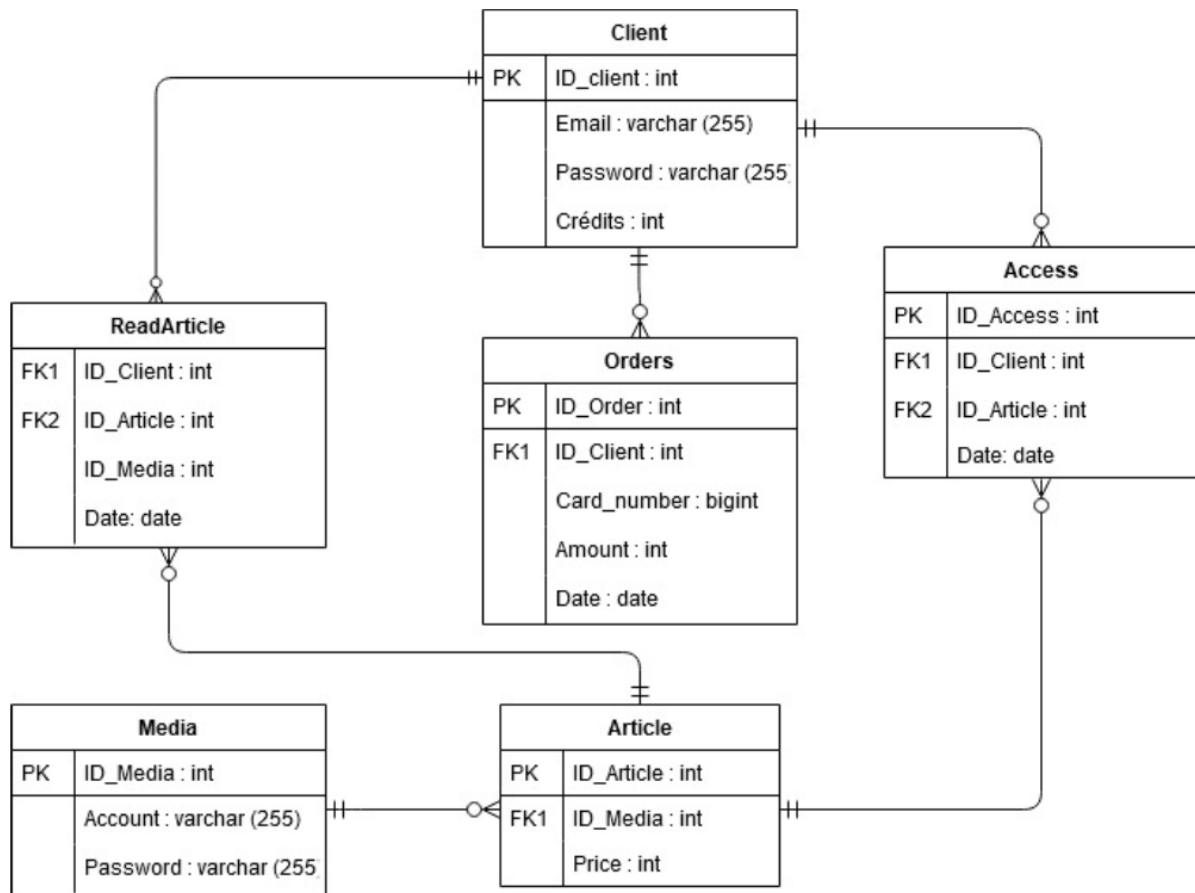


Figure 7 : Modèle entité-association

Un média possède un ou plusieurs articles sur sa plateforme, et un article est forcément lié à un fournisseur d'articles

Un article peut être associé à un ou plusieurs droits d'accès à celui-ci, mais un droit d'accès est obligatoirement lié à un article.

Chaque fois qu'un client lit un article, la table ReadArticle enregistre la date et va permettre ensuite de faire des analyses avec ces données.

Un client peut avoir un ou plusieurs droits d'accès, mais chaque droit d'accès n'est lié qu'à un seul client. Quand celui-ci recharge ses crédits, les informations concernant le paiement sont enregistrées dans la table Orders.

5.3. Force de l'architecture

Le projet MyPress possède ses propres comptes utilisateurs, gère lui-même les crédits, les achats, les accès, etc.

Cette indépendance présente plusieurs avantages :

- La sécurité de notre plateforme se concentre chez nous, celle-ci est peu influencée par le fonctionnement et la sécurité des sites des différents fournisseurs d'articles de presse
- Etant donné que tout est géré par la plateforme, le projet est plus facilement implémentable chez nos partenaires, même si la simplicité d'intégration dépend également des technologies utilisées par le fournisseur d'articles de presse
- Les informations concernant les accès et les achats sont enregistrées dans notre base de données. Il est alors possible de faire des statistiques afin de connaître les préférences des lecteurs, les périodes durant lesquelles ils n'achètent plus d'articles, les articles qui les intéressent le plus, etc.

6. Design du prototype

Une fois que le problème est modélisé, il faut maintenant réaliser un prototype qui peut le résoudre. Celui-ci est composé de différents éléments et choix technologiques qui sont faits lors de l'élaboration de l'architecture du projet MyPress, la sécurisation de celui-ci, et la réflexion sur la simplification du code afin de faciliter l'intégration du projet avec les autres médias partenaires.

6.1. But

Lors de ce travail de fin d'études, le but est de démontrer la faisabilité d'un projet (proof of concept) sans forcer de choix technologiques, afin de laisser aux futurs développeurs le choix sur les différentes possibilités de réalisation et d'implémentation.

Le prototype a pour but de démontrer que le projet est réalisable et fonctionnel, en proposant un exemple d'application qui répond à la problématique, à savoir permettre aux consommateurs de lire les articles de presse à la pièce dans les différents médias partenaires.

Celui-ci aura comme objectif de tester l'architecture proposée à travers différents scénarios de tests et de faire un déploiement du projet. De plus, il permettra de prévoir un audit de sécurité afin d'éviter certaines attaques informatiques.

6.2. Schéma relationnel du prototype

Comme nous l'avons vu plus haut, le choix du modèle de base de données s'est porté sur un modèle relationnel. Outre les raisons déjà évoquées, il simplifie certaines qui intéresseraient les différents fournisseurs d'article de presse partenaires du projet.

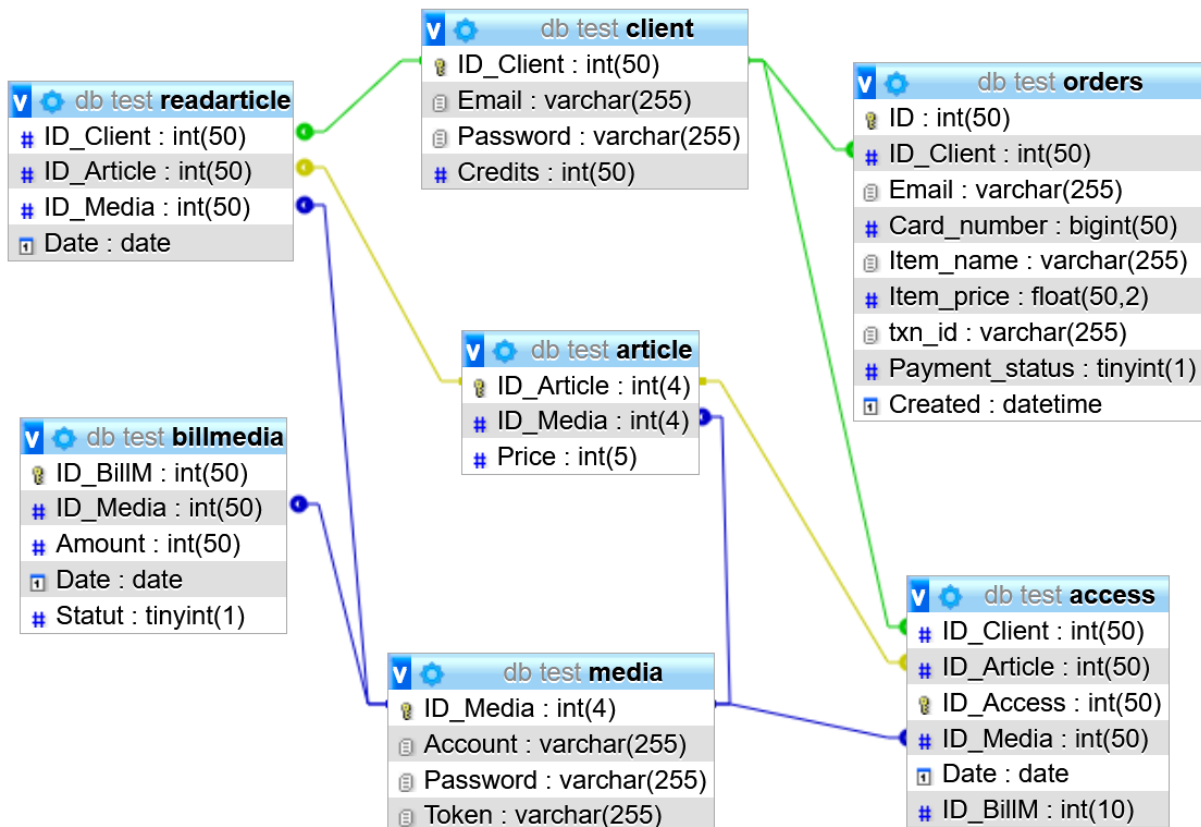


Figure 8 : Schéma relationnel de la base de données

Le prototype va utiliser ces tables, qui ont chacune un rôle :

- Table Article : stocke les articles accessibles grâce à la plateforme, ainsi que leur prix de vente en euros
- Table Client : stocke les données de tous les lecteurs
- Table Media : stocke les données de tous les médias partenaires
- Table Access : pour chaque client, stocke les articles achetés
- Table Orders : gère les informations concernant le rechargement de crédits

La colonne prix dans la table « Article » nous permet de gérer directement le prix d'un article, qui est fixé par le site de presse et l'équipe MyPress

Il aurait été possible de ne pas garder cette colonne, et d'inscrire le prix de l'article dans la requête émise par le client, lorsqu'il souhaite acheter un article. Ce mode de fonctionnement nous aurait permis de ne pas avoir à stocker dans notre base de données une nouvelle ligne, chaque fois qu'un fournisseur publie un article. Cependant, mettre le prix dans la requête est un risque que nous avons choisi d'éviter, étant donné que la requête se fait côté client et est donc émise par le consommateur. Celui-ci pourrait alors manipuler la requête afin de, par exemple, modifier le prix de l'article et ainsi avoir accès à de nombreux articles gratuitement.

Certaines tables permettront de faire des statistiques sur l'utilisation de la plateforme, ou de gérer le paiement des fournisseurs d'article de presse, etc. Il est possible de savoir quels sont les articles les plus achetés grâce à la table Access. On peut également savoir combien de fois un lecteur relit un article, et donc quels sont les articles qui font revenir un lecteur sur le site de presse, notamment grâce à la table ReadArticle qui stocke la date chaque fois qu'un client va lire un article.

Quelques problèmes subsistent. En effet, l'ID des lecteurs et médias est encodés comme un simple nombre, ce qui représente un grand risque puisqu'il est possible pour un pirate de modifier son cookie, et d'y inscrire l'ID de quelqu'un d'autre pour avoir accès à son compte. Il faudrait donc hasher l'ID des utilisateurs pour éviter qu'un pirate n'usurpe le compte d'un autre.

De plus, afin de récupérer les données concernant les articles des sites de presses, le développeur doit pour le moment encoder ses données à la main, mais il serait envisageable de créer une fonction avec l'API, où le site de presse envoie avec une requête les données de l'article. L'article serait alors enregistré automatiquement dans notre base de données et pourrait être consulté par un lecteur avec son compte MyPress.

Concernant les réglementations GDPR³, une convention a été préparée mais n'a pas été appliqué sur le prototype. Son impact modifiera la façon de stocker les données utilisateurs, tel que le nom de compte, l'adresse et le numéro de compte bancaire, stocker ses données pour une durée limitée, en plus de permettre à un utilisateur de supprimer son compte et modifier ses données personnelles, ce qui n'est actuellement pas possible.

³ Règlement général sur la protection des données

6.3. Architecture MVC

L'architecture du prototype suit une architecture répandue dans le domaine des applications Web : l'architecture Web-MVC⁴. Celle-ci se caractérise par une séparation des tâches en fonctions des rôles :

- Le contrôleur gère les requêtes utilisateurs et envoie les informations à la vue.
- La vue se charge d'afficher la réponse HTML, JavaScript, etc. à l'utilisateur.
- Le modèle se charge de créer et modifier les objets, et permet de dialoguer avec la base de données.

Dans le prototype présenté ici, le modèle est un modèle procédural. Il va contenir des fonctions permettant principalement d'accéder à la base de données, de récupérer des informations, d'en ajouter ou de modifier certaines données.

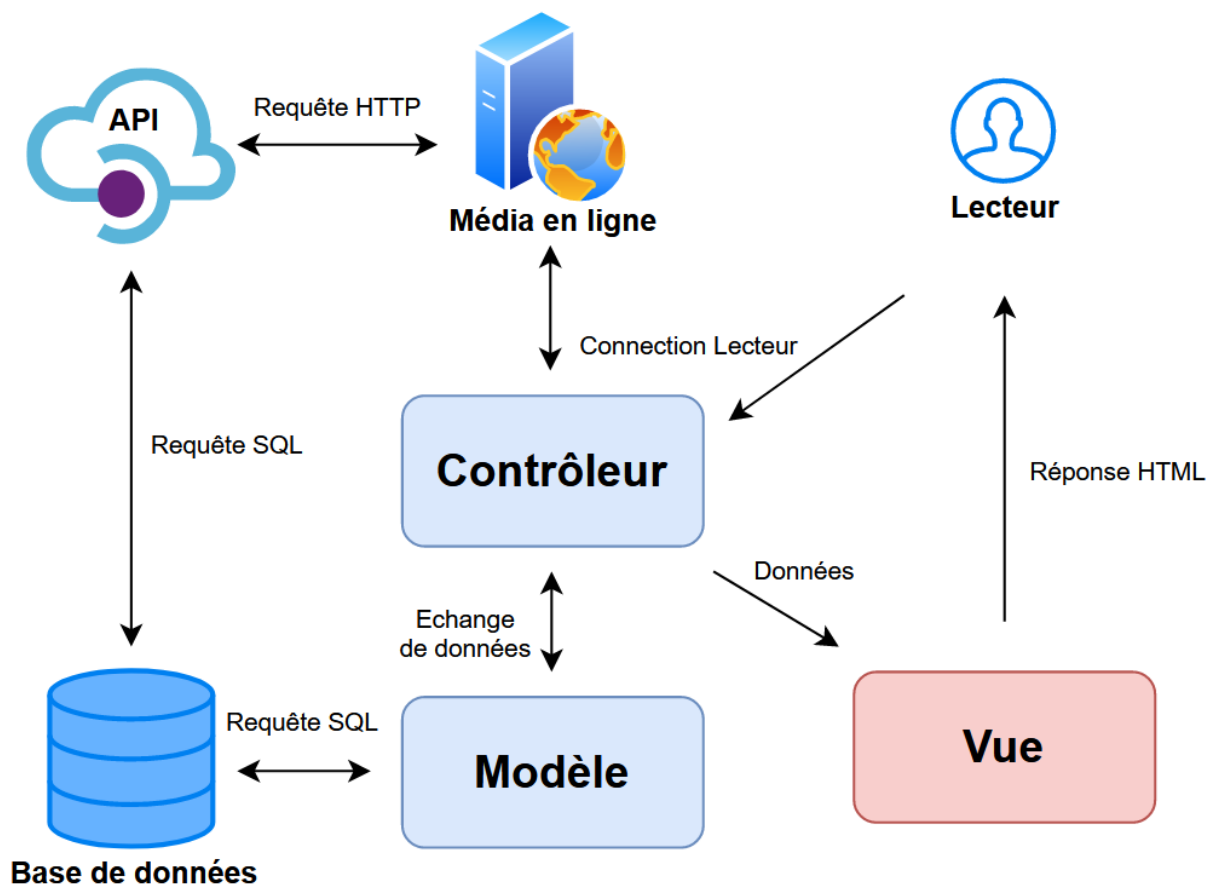


Figure 9 : Schéma de l'architecture du prototype

⁴ Architecture du type modèle Vue-Contrôleur

6.4. Analyse fonctionnelle

Une analyse fonctionnelle permet de déterminer quels seront les acteurs et les fonctionnalités d'un projet. Pour faire cette analyse, des diagrammes UML ont été réalisés.

Les diagrammes UML sont des schémas qui permettent de comprendre le projet en exposant son fonctionnement ainsi que les interactions qu'il y a entre les acteurs du projet, en montrant leurs actions. Ils synthétisent en un schéma les fonctionnalités d'un projet

6.4.1. Identification des fonctionnalités.

Ce diagramme de cas d'utilisations décrit les différentes fonctionnalités qui seront proposées aux utilisateurs et acteurs du projet ainsi que les interactions entre eux

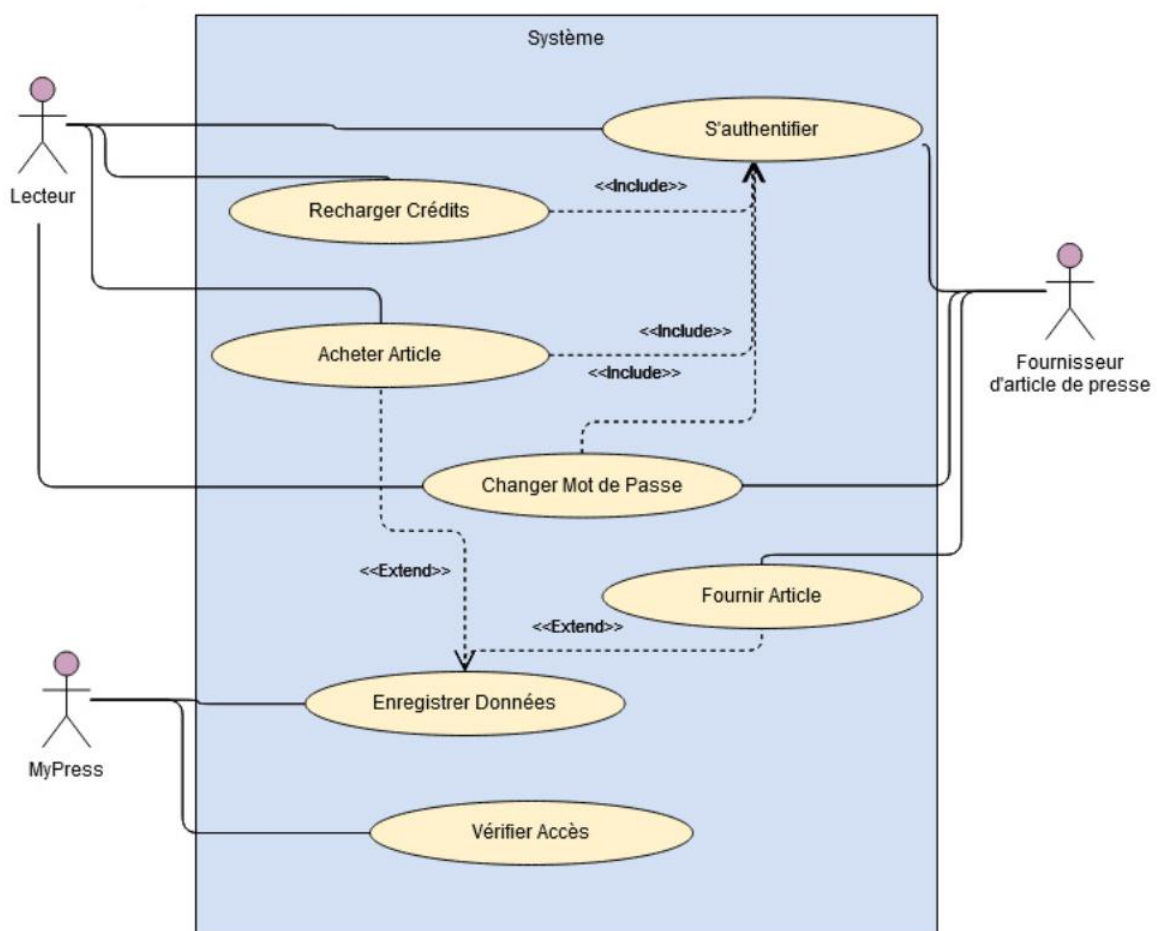


Figure 10 : Diagramme cas d'utilisation

Ce sont les fonctionnalités de bases qu'il faut développer pour le prototype pour qu'il puisse réaliser la gestion des achats d'articles et la gestion des accès aux articles

6.4.2. Gestion des connexions et des articles

Ce diagramme d'activité décrit les étapes parcourues par un consommateur lorsqu'il souhaite se connecter sur un site de presse ou sur notre plateforme.

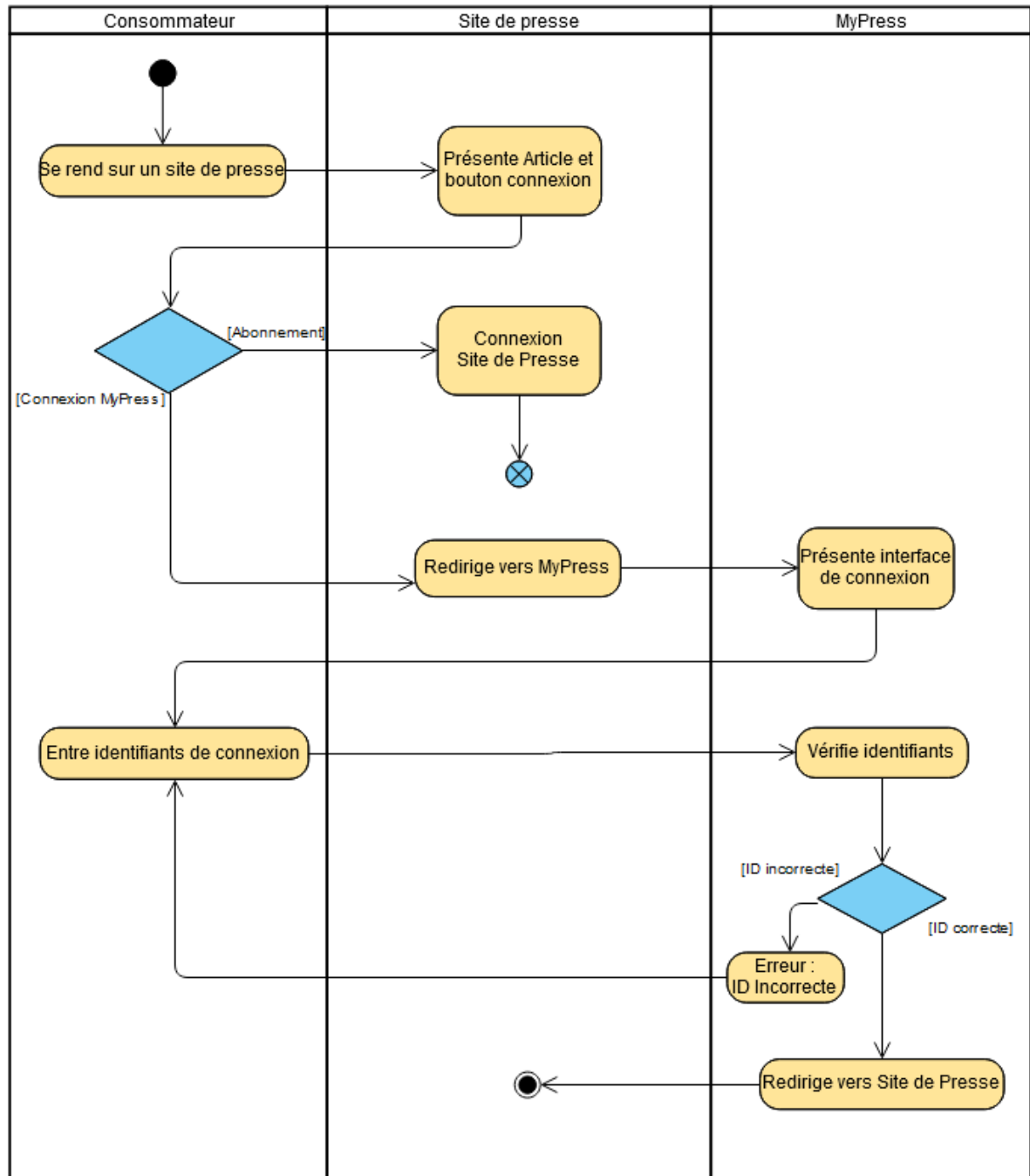


Figure 11 : Diagramme d'activité - Connexion MyPress

Lorsqu'un lecteur souhaite lire un article, il peut soit se connecter à son abonnement classique, ou se connecter à son compte MyPress, en passant par notre plateforme afin de s'identifier. Il sera ensuite redirigé vers le site de presse.

Ce diagramme décrit les étapes concernant l'accès à un article du site de presse. Le client, une fois connecté, tentera de lire le contenu d'un article, et celui-ci lui sera affiché uniquement s'il a acheté cet article.

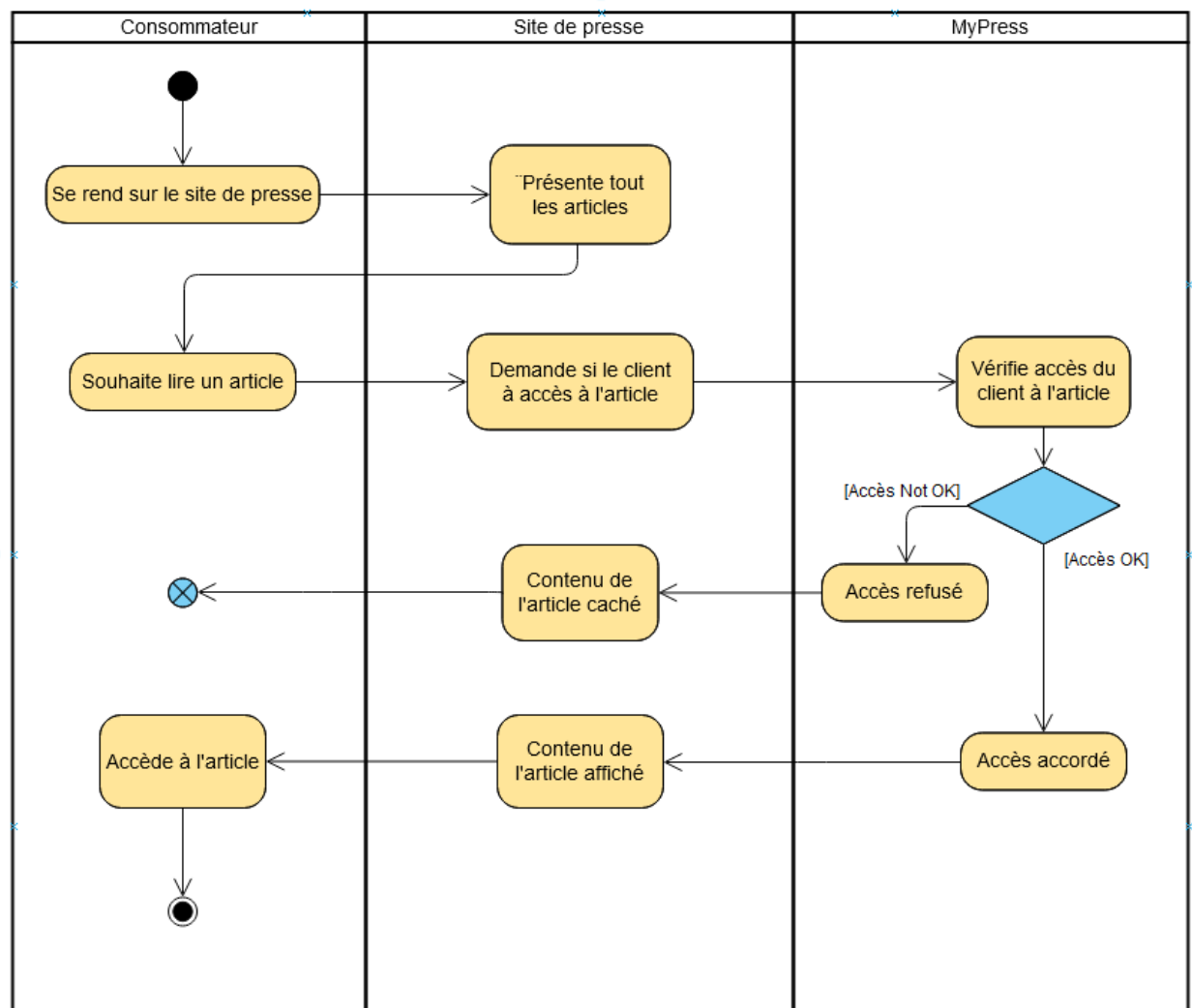


Figure 12 : Diagramme d'activité - Accès à l'article

Lorsqu'un lecteur souhaite lire un article, le site de presse doit vérifier si ce client a le droit de le lire, il contactera alors MyPress, qui fera cette vérification et ensuite répondra au site de presse que l'accès est accordé ou refusé

Ce diagramme représente les étapes à propos de l'achat d'un nouvel article par un consommateur.

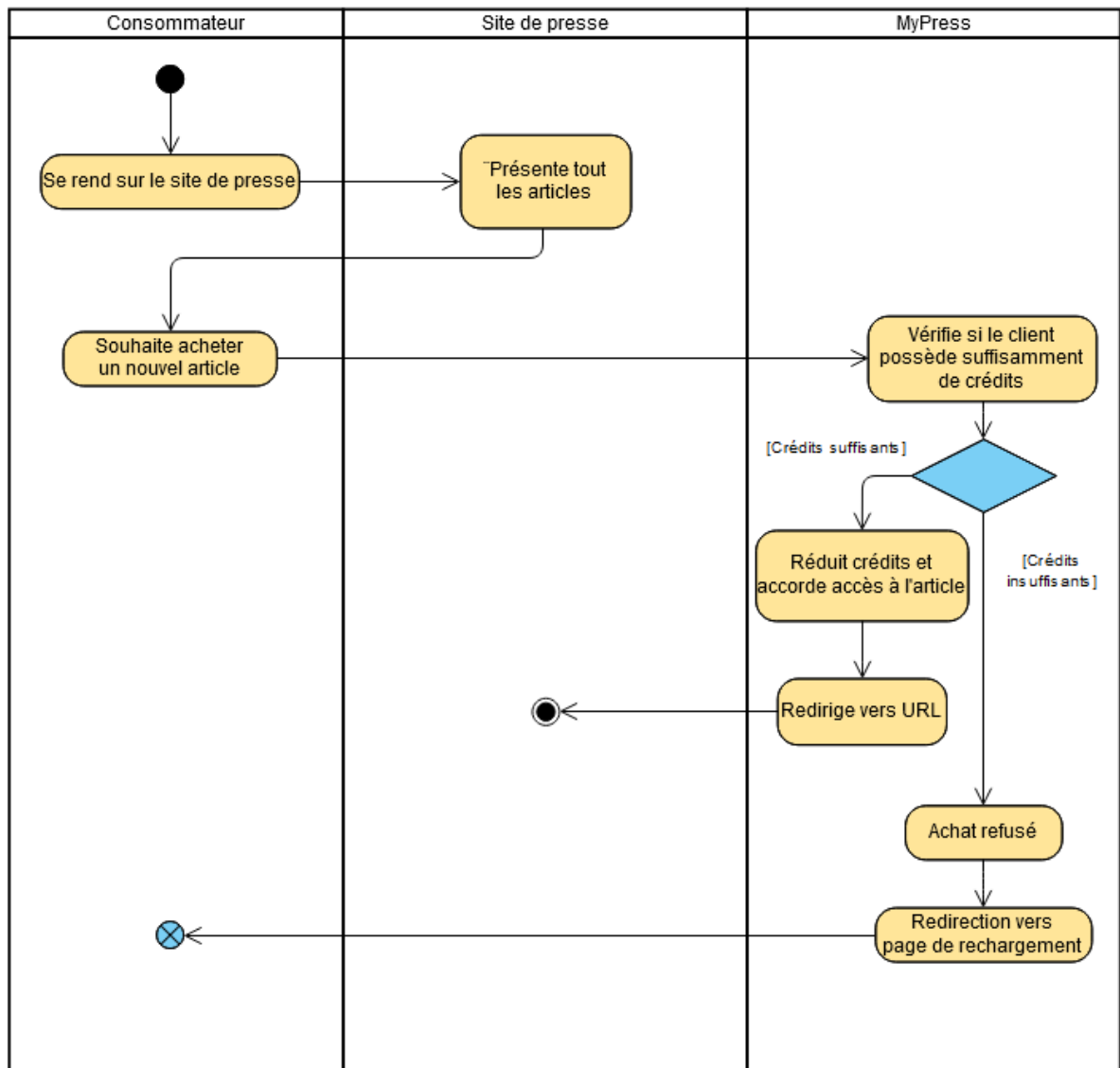


Figure 13 : Diagramme d'activité - Achat d'un nouvel article

L'achat d'un article se fera grâce à un bouton JavaScript, qui effectue une redirection vers notre plateforme afin de procéder à l'achat. Le site de presse ne fait que présenter le bouton JavaScript au lecteur, le reste de la procédure est géré par la plateforme MyPress.

6.4.3. Token de sécurité API

Ce diagramme de séquence représente le processus par lequel va passer un fournisseur d'article de presse pour récupérer son token de sécurité.

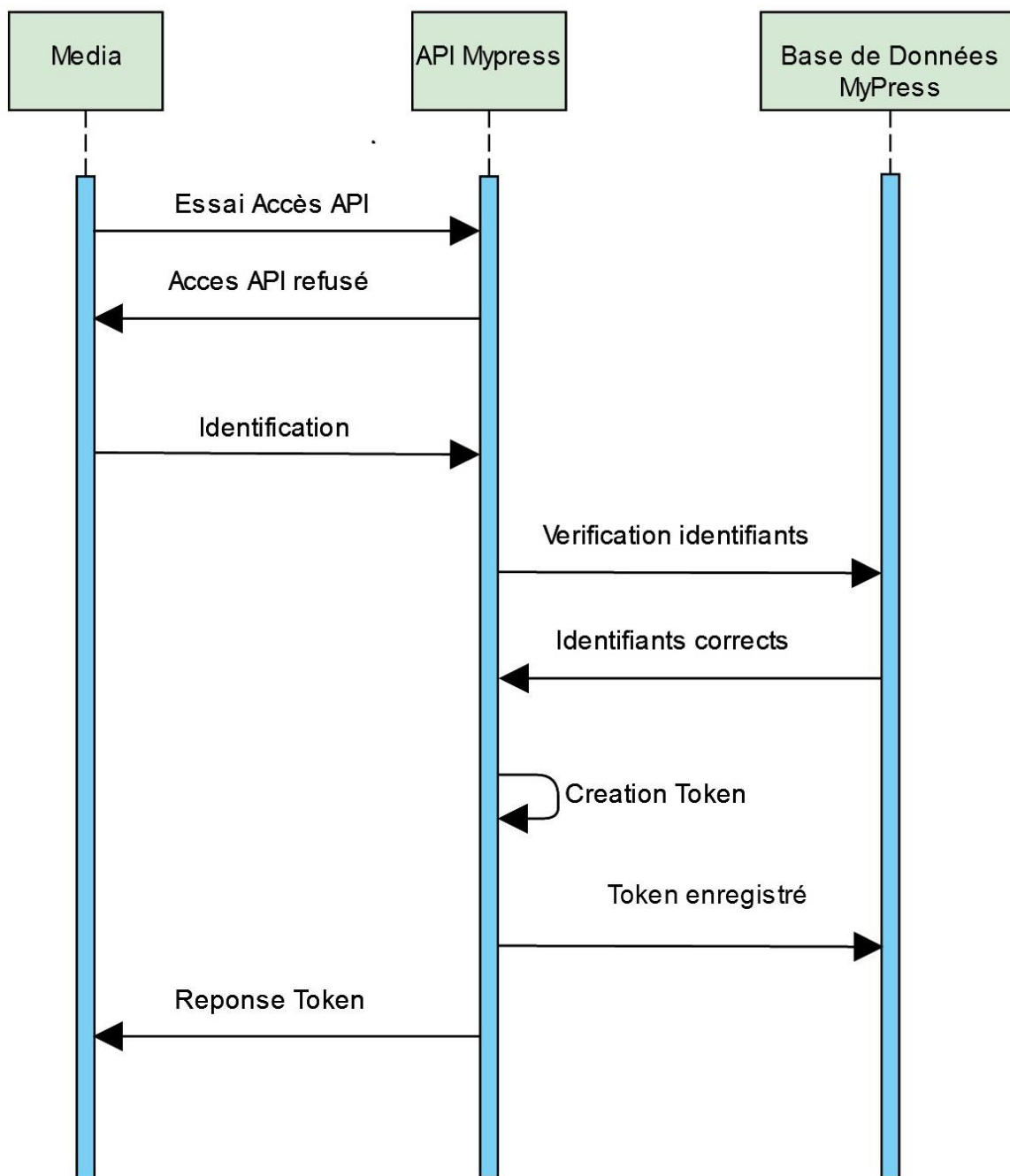


Figure 14 : Diagramme de séquence - Obtention du Token d'API

L'obtention du token est nécessaire pour toutes utilisations de l'API, excepté pour s'identifier auprès de celle-ci.

6.5. Transfert des données et Cookie

Le prototype MyPress étant une plateforme intermédiaire permettant à un client d'accéder à un contenu d'un autre site, il y a certaines données qui sont transférées entre les acteurs du projet.

Les principales informations qui transitent entre ceux-ci sont les données permettant d'identifier un site de presse et un lecteur, ainsi que les données permettant de gérer les accès aux articles du site de presse.

6.5.1. Connexion du média

Avant toute chose, le site de presse doit se connecter à la plateforme afin de recevoir son token lui permettant de faire des requêtes avec l'API de MyPress.

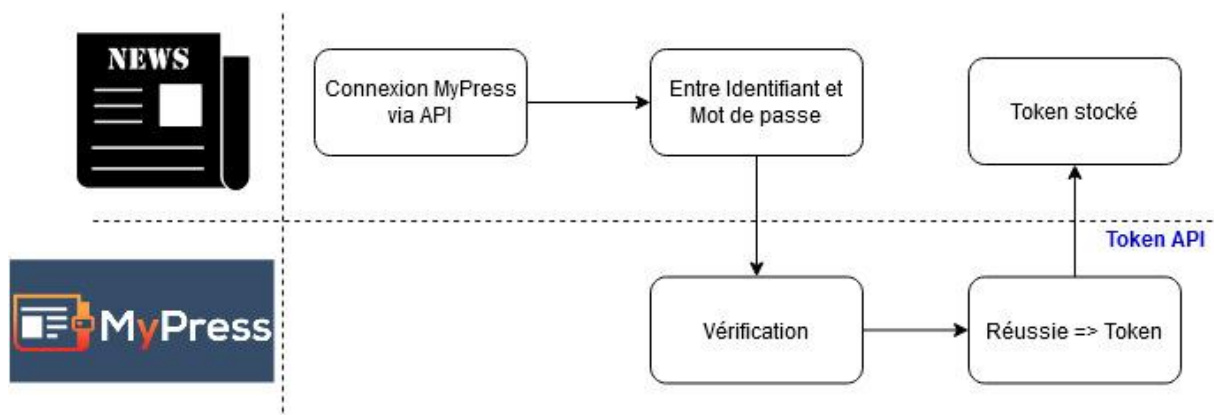


Figure 15 : Connexion du site de presse

Afin de s'authentifier, un média doit envoyer ses coordonnées de connexion via l'API. Une fois vérifié, l'API lui renverra un token et l'enregistrera dans la base de données MyPress. Si un média possède un token qui a expiré, celui-ci doit repasser par l'API afin de récupérer un nouveau token.

Ce token a pour but d'éviter l'accès à l'API par n'importe quel site ou personne. Le token est vérifié à chaque requête envoyée à l'API. Si le token ne correspond pas à celui inscrit dans la base de données de MyPress, la requête est rejetée.

6.5.2. Connexion utilisateur

Le client va se connecter à son compte MyPress directement sur le site de la plateforme, qui sera utilisé afin de savoir s'il a le droit de lire un article.

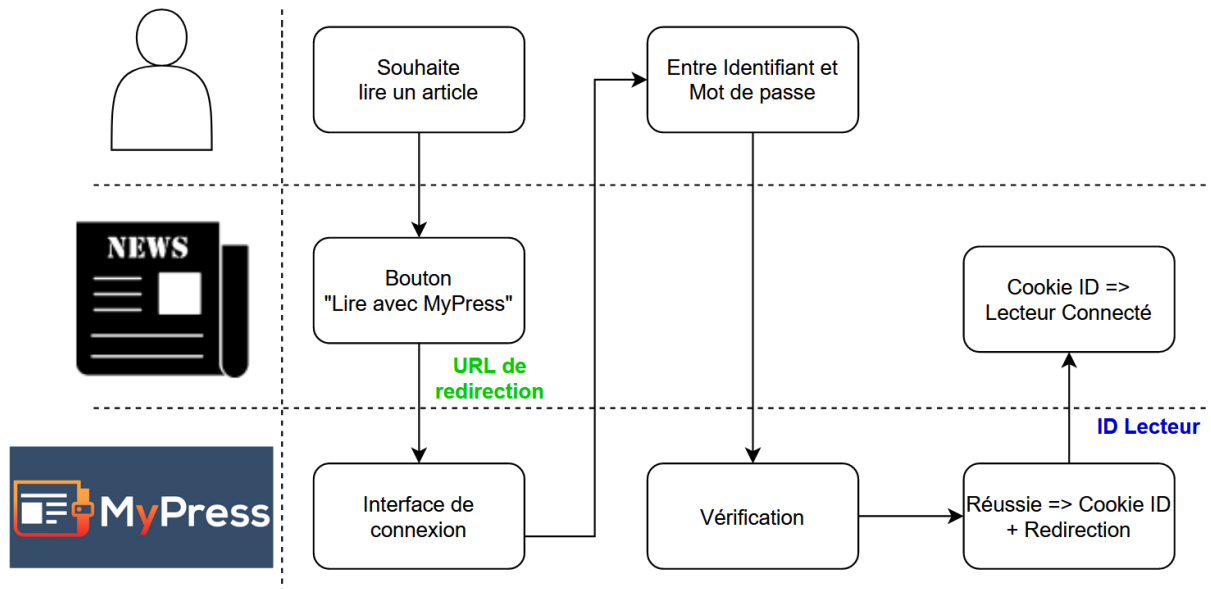


Figure 16 : Connexion utilisateur

Le client, lorsqu'il cliquera sur le bouton « Lire avec MyPress » du site de presse, sera redirigé vers le site de la plateforme afin qu'il puisse se connecter à son compte MyPress avec ses coordonnées de connexion. Une fois connecté, l'ID du client est retourné afin qu'il puisse être stocké dans un cookie, et le client est redirigé vers le site de presse.

6.5.3. Vérification des accès aux articles

Le site de presse doit faire une requête à l'API de MyPress afin de déterminer si un client connecté à son compte MyPress a le droit de lire le contenu d'un de ses articles.

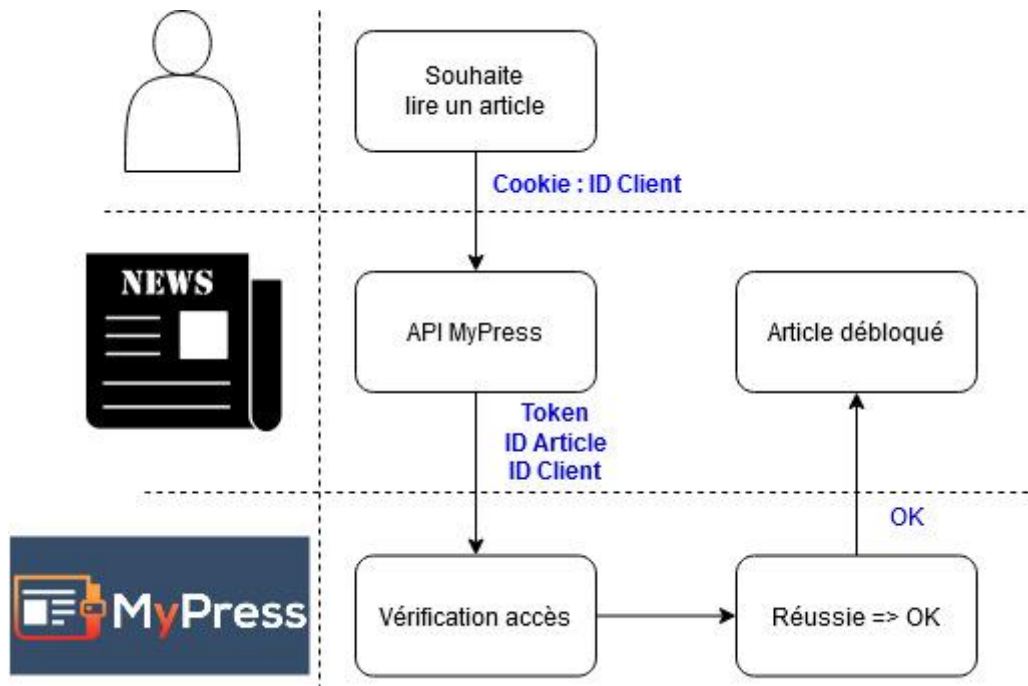


Figure 17 : Vérification des accès

Le média inscrit dans cette requête quelques données permettant à l'API de procéder à la vérification du droit d'accès :

- Token de sécurité, qui vérifie si le site de presse a le droit de faire des requêtes à l'API
- ID du lecteur, récupéré dans le cookie du site de presse
- ID de l'article

Une fois la vérification effectuée du côté de MyPress, la plateforme répond à la requête par un « OK » si la requête a réussi, ou par une erreur si la requête a échoué.

Plusieurs erreurs peuvent se produire et faire échouer la requête. Les plus probables sont :

- Le client ne possède pas le droit d'accès, et donc la vérification d'accès envoie un message d'erreur : False

- Le site de presse ne possède pas un token valide lui permettant de faire des requêtes avec l'API. Celle-ci va donc directement lui refuser la requête et lui envoyer un message d'erreur : Requête non autorisée, token incorrect

6.5.4. Achat d'un article

Un lecteur connecté à son compte peut acheter un article afin de lire son contenu grâce à un bouton codé en JavaScript sur le site de presse. Celui-ci mène l'utilisateur sur la plateforme MyPress afin d'effectuer l'achat. Le client doit avoir assez de crédits afin de faire cet achat. Dans le cas contraire, l'achat sera annulé, un message en informera le client et il sera redirigé vers la page de rechargement de crédits.

Afin de procéder à l'achat, les données suivantes sont envoyées à la plateforme :

- ID du lecteur
- ID de l'article
- ID du provider
- URL de retour, afin de rediriger le client sur la page de l'article acheté une fois la transaction effectuée.

6.6. Authentification

Afin d'utiliser les services proposés par la plateforme, les utilisateurs doivent s'inscrire et ensuite se connecter à leur compte MyPress. À l'inscription, ils entrent leur adresse mail qui sera utilisée en tant qu'identifiant, ainsi qu'un mot de passe afin de garantir la protection des données du compte.

Une fois connecté, le lecteur a accès à son compte, sur lequel il peut voir quels sont les articles qu'il a achetés et le nombre de crédits qu'il lui reste. Il a également accès à l'historique de ses précédents achats.

Une page permet également à l'utilisateur de changer son mot de passe quand il le souhaite.

Les médias partenaires du projet possèdent également un compte sur la plateforme MyPress. Ce compte leur permet de voir quelques informations, telles que le nombre de fois où un article de leur site a été acheté à travers notre plateforme.

Les mots de passe des lecteurs et des médias partenaires sont stockés dans la base de données du projet et sont hashés suivant l'algorithme de hashage Blowfish. Une section du chapitre « Sécurité » explique cet algorithme de hashage.

6.7. Rechargement de crédits

Les lecteurs payent l'achat d'articles en utilisant une monnaie virtuelle propre à la plateforme, sous forme de crédits. Lorsqu'ils sont à court de crédits, ils rechargent leur compte en inscrivant un numéro de carte bancaire, date d'expiration de la carte, etc.

La gestion des paiements est faite grâce à l'utilisation de Stripe, une solution de traitement de paiements en ligne.

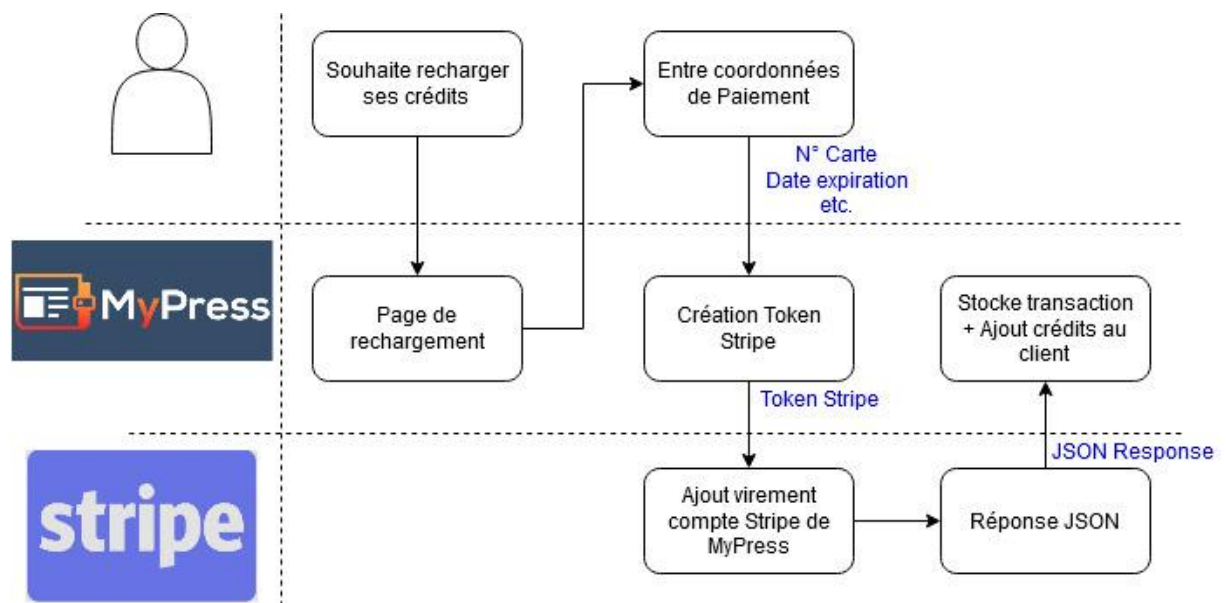


Figure 18 : Rechargement des crédits

Ainsi, les informations de paiements sont transférées à la suite d'outils Stripe, qui va effectuer les virements bancaires du consommateur vers le compte Stripe de la plateforme MyPress. Une fois effectué, Stripe nous renvoie un JSON contenant des informations de paiement, telles que le nom du client, le montant, la date de création du paiement, etc.

L'utilisation de Stripe simplifie l'implémentation d'un processus de paiement, mais celui-ci prend une commission à chaque paiement effectué par un consommateur. Dès lors, nous ne recevons pas la totalité de l'argent versé par le consommateur lorsqu'il achète des crédits, mais cet outil suffit pour la création d'un prototype.

7. Test

Après avoir créé un prototype fonctionnel, des tests sont faits sur celui-ci afin de vérifier son fonctionnement à travers différents scénarios d'utilisations possibles. L'API du projet a également été testée afin de vérifier sa sécurité et sa réponse aux différentes requêtes.

7.1. Stratégie

Afin de tester les fonctionnalités de la plateforme, deux cas ont été préparés.

D'une part, un site de presse fictif a été mis en place. Une fois créé, il faut ensuite le faire dialoguer avec la plateforme MyPress, et ainsi vérifier son bon fonctionnement.

D'autre part, l'outil a été testé sur un site externe, programmé par une autre personne.

7.2. Démarche de test

Dans le but de tester notre projet, différents scénarios d'utilisation de la plateforme ont été imaginés, que ce soit pour les lecteurs ou les fournisseurs d'articles de presse :

- Un client déconnecté tente de lire un article
→ Résultat attendue : Il doit se connecter
- Un client connecté tente de lire un article qu'il n'a pas acheté
→ Résultat attendue : Le contenu de l'article lui est caché
- Un client tente d'acheter un article sans crédits MyPress
→ Résultat attendue : Achat annulé et redirection vers la page de rechargement
- Un client avec suffisamment de crédits MyPress achète un article
→ Résultat attendue : L'achat est enregistré et les crédits du client sont débités
- Un client lit un article qu'il a déjà acheté auparavant
→ Résultat attendue : Le contenu de l'article lui est disponible
- Un client change de mot de passe
→ Résultat attendue : La connexion à son compte demande maintenant le nouveau mot de passe

- Un média qui ne possède pas le token de sécurité tente de faire une requête à l'API
→ Résultat attendue : La requête est rejetée et un message d'erreur « Non autorisé » est envoyé
- Un média qui possède un token de sécurité invalide tente de faire une requête à l'API
→ Résultat attendue : La requête est rejetée et un message d'erreur « Non autorisé » est envoyé
- Une personne tente de récupérer un token en entrant de mauvais identifiants
→ Résultat attendu : Accès refusé

Une fois ces scénarios de test établis, le projet est testé en suivant une boucle telle que celle du Test-Driven-Development :

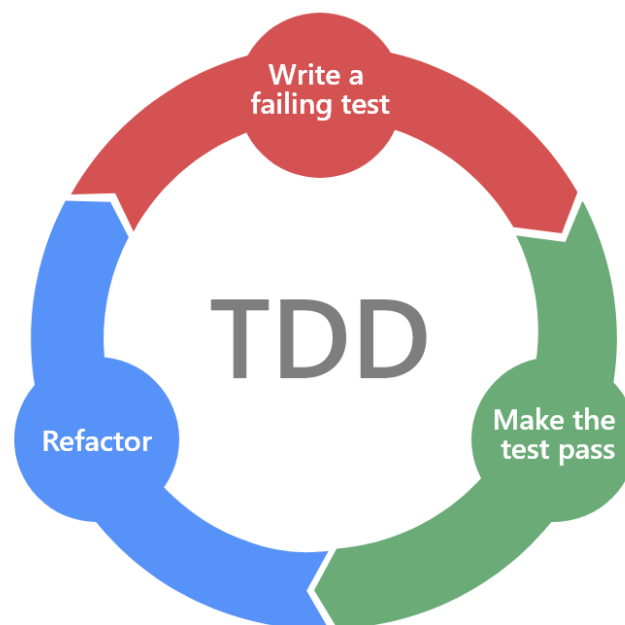


Figure 19 : Boucle de test TDD

Si le test échoue, et donc le résultat obtenu est différent du résultat attendu (une autre page s'ouvre, un nouvel accès est accordé alors que le client n'a pas de crédits, etc.), un debugging permet de modifier le code. Ensuite, le scénario de test est à nouveau appliqué afin de vérifier que le test passe, et enfin on simplifie le code.

Si le test passe, c'est-à-dire si le résultat attendu est le même que le résultat obtenu, le test est considéré comme réussi et on passe au test suivant.

7.3. Site fictif : BelgiumPost

BelgiumPost est un « faux » site de presse fonctionnel, capable de montrer ou cacher le contenu d'un article aux clients. Il va permettre de voir comment il faut interconnecter MyPress avec un site de presse et de mettre en lumière les problèmes de la plateforme afin de les corriger.

Ce site présente des articles gratuits et des articles payants, qui pourront être lus grâce à un compte sur la plateforme MyPress. Lorsque l'on clique sur un article payant, le contenu de l'article est caché, et un bouton permet de le débloquent grâce à MyPress. Le bouton redirige le lecteur sur la page de connexion de MyPress. Une fois connecté, il est redirigé vers une URL choisie par le site de presse.

Ainsi, le consommateur est de retour sur le site de presse et connecté à MyPress. Cependant, le contenu de l'article ne lui est toujours pas dévoilé, étant donné qu'il doit encore acheter l'article en question. Un bouton permet donc d'acheter l'article et de débiter le compte MyPress de l'utilisateur pour cet achat. L'article est alors débloquent.

Ce site a pour but de tester la plateforme et d'évaluer si les scripts qu'il faut rajouter au site de presse peuvent être intégrés sur un nouveau site

Il faut d'abord identifier les morceaux de code nécessaire à cette coopération :

- Code de la requête permettant au site de presse de recevoir son token et d'avoir accès à l'API de MyPress
- Code de la requête à l'API afin de vérifier si un lecteur a accès à un article ou non
- Code contenant le bouton permettant d'acheter un article en utilisant son compte MyPress

Pour tester le prototype MyPress avec ce site de presse, il faut interconnecter les deux plateformes, et enfin, les scénarios de tests sont lancés les uns après les autres.

7.4. Site externe

Un second site de test a été créé afin de tester l'intégration du code à importer et permettre la coopération entre le site de presse et la plateforme MyPress. Ce second site a été créé par un autre développeur, ce qui permet d'éviter tout type d'influence qui pourrait permettre un dialogue plus facile avec la plateforme MyPress. On a ainsi pu tester l'outil avec un site complètement extérieur au projet.

Une fois interconnecté grâce aux mêmes morceaux de code que ceux utilisés pour BelgiumPost, ce second site a permis de tester le fonctionnement du site et de mettre en lumière quelques problèmes, tels que les soucis liés au format des variables inscrites dans la requête d'achat d'article, etc..

Pour ce nouveau site, les tests cités précédemment ont également été lancés, et d'autres scénarios ont été testés :

- Achat d'un article qui n'est pas enregistré dans la base de données MyPress
- Utilisation d'un même compte client sur deux sites de presse différents (mise à jour des crédits, accès des articles, etc.)
- Accès à un article d'un même ID, mais d'un fournisseur de presse différent

Ces tests ont permis de résoudre des problèmes d'interconnexion entre notre projet et un site extérieur, et ont permis de simplifier l'intégration du projet MyPress au sein d'un site de presse, extérieur au projet.

8. Sécurité

Cette section explique différents risques et dangers auxquels il faut prêter attention afin d'assurer le bon fonctionnement du projet.

Elle montre également les protections mises en place afin de se protéger de certaines attaques communes lorsqu'une plateforme est mise en ligne

8.1. Attaque et danger

La mise en place d'une plateforme en ligne possède son lot de risques et de failles qui pourraient permettre à un utilisateur malveillant de pirater le système et d'accéder à des données sensibles ou de modifier le comportement de la plateforme.

Il existe de nombreuses attaques utilisées par les hackers, visant ce genre de plateforme, telles que :

- **Attaque DDOS** : Attaque consistant à envoyer un nombre très élevé de requêtes par seconde à une plateforme en ligne afin de la surcharger et que celle-ci cesse de fonctionner ou que le niveau de service qu'elle rend se dégrade fortement.
- **Attaque XSS** : Attaque permettant d'entrer du code dans un champ, qui sera ensuite exécuté par le site sur les ordinateurs des clients qui naviguent dessus.
- **Injection SQL** : Injection permettant de modifier une requête SQL afin de récupérer des données ou de modifier l'état de la base de données d'un site web
- **Sniff de paquet** : Attaque consistant à voler un paquet de données envoyé sur le réseau afin d'avoir accès aux informations qu'il contient

Les principaux composants à protéger sont le serveur, l'API et la base de données du projet

Une protection physique du système est aussi à prendre en compte, telle que la garde des backups de la base de données dans un autre lieu physique, pour éviter qu'en cas de crash du système ou de corruption de la base de données, ces informations ne soient définitivement perdues.

De plus, la particularité de notre plateforme, qui dialogue avec les sites de presse afin de donner accès à un contenu sur le site de presse, fait qu'il est très important de veiller à sécuriser les informations de paiements, le fonctionnement de l'achat de crédits, la connexion avec le site de presse, etc. afin d'éviter que l'intégration de notre plateforme sur les sites partenaires ne créer de nouvelles failles pour ceux-ci.

8.2. Analyse de risque

Une analyse de risque permet de déterminer la dangerosité d'une attaque sur un composant du projet en évaluant différentes caractéristiques de celui-ci et va ensuite estimer le niveau de dangerosité ou l'éventualité de cette attaque.

Les caractéristiques analysées ici sont :

- Bien : l'objet du système qui va subir l'attaque
- Menace / Vulnérabilité : le type d'attaque
- Contrôle existant : les éléments déjà mis en place qui permettent de protéger le bien de l'attaque
- Probabilité : la probabilité qu'une attaque de ce type se produise sur le projet

Ici, les probabilités sont classées sur une échelle d'un à cinq :

Valeur	Probabilité	Définition
1	Rare	Peut se produire dans des circonstances exceptionnelles
2	Peu probable	Peut se produire à un moment donné, mais non prévu compte tenu des événements récents
3	Possible	Peut se produire à un moment, difficile à contrôler l'occurrence, facteurs extérieurs
4	Probable	Va probablement se produire, sans surprises
5	Presque certain	Prévu de se passer tôt ou tard, dans la plupart des circonstances

- Les conséquences sur l'état du projet, classées ici sur une échelle d'un à six :

Valeur	Conséquences	Définition
1	Insignifiant	<ul style="list-style-type: none"> • Vient d'une faille mineure dans une seule zone • Impact faible • corrigé en quelques jours
2	Mineur	<ul style="list-style-type: none"> • Faille dans une ou deux zones • Impact faible, influe sur une partie du projet • Corriger en moins d'une semaine
3	Modéré	<ul style="list-style-type: none"> • Faille de sécurité limitée • Impact jusqu'à deux semaines • Quelques coûts de réparation • Clients sont avertis du problème
4	Majeur	<ul style="list-style-type: none"> • Faille de sécurité • Impact de quatre à huit semaines • Management et ressources nécessaires pour la réparation • Client averti • Possibilité de perte business
5	Catastrophique	<ul style="list-style-type: none"> • Faille de sécurité majeure • Impact de plus de trois mois • Coûts de réparation élevés • Probable perte de client • Perte de confiance en l'entreprise
6	Apocalypse	<ul style="list-style-type: none"> • Plusieurs failles de sécurité majeures • Impact à durée indéterminé • Procédure judiciaire • Perte de business • Liquidation de l'entreprise probable

Pour le prototype réalisé ici, les analyses de risques vont se limiter au serveur du projet qui se charge de faire tourner la plateforme en ligne et à la base de données du prototype qui contient certaines données sensibles.

Ces analyses ne se concentrent ici que sur quelques attaques, celles qui ciblent les données importantes et qui empêchent le bon fonctionnement de la plateforme

8.2.1. Analyse de risque : Base de données MyPress

- Bien : Base de données MyPress
- Menace : Injection SQL
- Contrôle existant : Préparation des requêtes SQL, Hashage des mots de passe avec algorithme Blowfish
- Probabilité : Probable
 - ➔ L'injection SQL est une attaque très connue et souvent utilisée par les hackers, et la base de données MyPress contient des données sensibles qui pourraient intéresser les hackers
- Conséquences : Majeur
 - ➔ Cette faille de sécurité aura un impact non négligeable et la réparation de la base de données vers un état correct prendra quelque temps.
La plus grosse conséquence étant la perte du business des fournisseurs d'articles de presse partenaire

Ce type d'attaque possède un risque extrêmement haut, puisque les conséquences sont importantes et que la probabilité que ce type d'attaque arrive sur notre plateforme est élevée.

De plus, nous enregistrons des données sensibles sur les clients, mais aussi sur les fournisseurs d'articles de presse partenaire, telles que leur token d'accès à l'API, ou encore des informations concernant les échanges financiers entre ceux-ci et la plateforme MyPress.

Une attaque de ce type peut permettre à un hacker d'accorder frauduleusement des accès à des articles, ou de directement modifier le nombre de crédits des clients. Nous serions alors une faille permettant à des personnes de lire gratuitement les articles des médias partenaires.

Ce type d'attaque permettrait également à un hacker de récupérer les identifiants et mots de passe des lecteurs ou des médias partenaires. Cependant, les mots de passe étant hashés par la plateforme, il ne récupérera qu'un hash, ce qui ne lui permet pas de se connecter à un compte car celui-ci sera alors ré-hashé par la plateforme lors de la tentative de connexion.

8.2.2. Analyse de risque : API MyPress

- Asset : API MyPress
- Menace : Attaque DDoS
- Contrôle existant : JWT token d'accès à l'API
- Probabilité : Possible
 - ➔ L'attaque DDoS est très efficace lorsqu'une personne malintentionnée souhaite rendre inutilisable un service en ligne
- Conséquences : Modéré
 - ➔ Une attaque DDoS sur l'API empêcherait celle-ci de fonctionner et donc empêcherait les fournisseurs d'articles de presse de vérifier les accès afin de savoir quel client peut lire quel article.

Ce type d'attaque possède un haut risque, puisque la probabilité que ce type d'attaque arrive sur notre plateforme est élevée, mais les conséquences ne bloqueraient le fonctionnement que d'une partie de la plateforme

Une attaque DDoS n'apporte pas grand-chose aux hackers, mais elle empêche la plateforme de fonctionner correctement, ce qui peut distraire les agents chargés de la sécurité informatique en vue d'une autre attaque.

Afin de s'en protéger, il serait possible de limiter le nombre de requêtes maximum reçues dans un temps donné.

8.3. Certificat HTTPS

HTTP est un protocole de transmission d'informations client-serveur classique. La forme HTTPS est une amélioration sécurisée grâce à une couche de chiffrement, telle que le SSL.

Ce protocole sécurisé permet de vérifier un site web via un certificat d'authentification émis par une autorité réputée fiable. Il est couramment utilisé lors de transactions financières, ou sur les réseaux sociaux, afin de protéger la connexion et la vie privée d'un utilisateur.

Dans notre prototype, le HTTPS a pour but de protéger les comptes des utilisateurs et la confidentialité de leurs informations bancaires. Il est garanti par l'autorité Let's Encrypt.

8.4. Attaque XSS

Lorsqu'un champ d'un formulaire est proposé à un utilisateur, celui-ci peut l'utiliser pour y entrer une donnée, telle qu'une image suivie d'une description. Cependant, l'utilisateur peut également entrer, au lieu d'une description, du code HTML/JavaScript qui sera alors exécuté par le navigateur.

Il existe deux types d'attaques XSS :

- Non persistante : Le code injecté n'est pas sauvegardé par le serveur, il est par exemple directement affiché, comme pour un formulaire de recherche
- Persistante : Le code est sauvegardé par le serveur, lorsqu'un utilisateur upload une image avec une description contenant du code sur un serveur.

Lors d'une attaque persistante, étant donné que le code injecté est enregistré sur un serveur, il est possible de récupérer les données confidentielles d'un autre utilisateur si celui-ci clique sur le lien contenant l'image et la description frauduleuse, avec le code JavaScript qui sera interprété par le navigateur de la victime.

De cette façon, il est par exemple possible de récupérer les cookies d'un autre utilisateur.

Afin de se protéger des attaques XSS, il faut sécuriser toutes les entrées et formulaires qui pourraient permettre à une personne malintentionnée d'injecter du code.

La fonction « htmlspecialchars » permet d'ignorer un code entré dans un formulaire et de le considérer directement comme une chaîne de caractère, afin de ne pas interpréter le code injecté dans ce formulaire

Ainsi, si un utilisateur entre dans la case Nom : `Jean` il sera traité comme étant `Jean`, et non comme étant **Jean**

8.5. Injection SQL

Les opérations effectuées afin de récupérer une donnée, de modifier ou d'insérer une nouvelle information dans la base de données se font à travers des requêtes SQL.

Celles-ci permettent de communiquer avec la base de données avec un type de requête (INSERT pour insérer une nouvelle ligne de donnée) ainsi qu'un corps de requêtes, qui va contenir les données que l'on souhaite ajouter dans la base de données

L'injection SQL est un type d'attaque permettant de modifier une requête afin d'insérer de fausses données ou de récupérer des données qui sont normalement confidentielles. En effet, de nombreuses requêtes demandent aux utilisateurs d'entrer des informations afin d'effectuer la requête.

Exemple :

Requête SQL : `SELECT * FROM admin WHERE login = X AND password = X`

L'utilisateur doit normalement entrer à la place du X la donnée demandée, telle que le login de l'admin.

⇒ Requête SQL : `SELECT * FROM admin WHERE login = administrator AND password = mot_de_passe`

Mais si l'utilisateur est malintentionné et que celui-ci, au lieu d'insérer une donnée dans le formulaire, insère du code SQL, celui-ci sera alors inscrit dans la requête et la modifiera lors de son exécution.

⇒ Requête SQL : `SELECT * FROM admin WHERE login = administrator AND password = ' OR 1 = 1`

Ainsi, la partie concernant le password va être ignorée à cause de l'injection du code SQL «OR 1=1 », qui est toujours vraie. La requête sera donc effectuée sans que l'utilisateur entre le mot de passe de l'administrateur.

Afin de se protéger de ce type d'attaque, il faut préparer la requête afin que les paramètres de la requête soient validés avant que l'utilisateur n'entre ses données. Il n'est donc plus possible d'ignorer les paramètres de la requête.

Ce type d'attaque est dangereux dans plusieurs requêtes de notre prototype, notamment pour les requêtes de connexions, où un utilisateur pourrait accéder au compte d'un autre utilisateur ou d'un média de presse.

Également, lors de l'achat d'un article, un utilisateur pourrait manipuler la requête afin d'acheter tous les articles qu'il souhaite sans payer leur prix. Notre prototype serait alors une faille de sécurité permettant aux lecteurs de lire les articles de sites de presse gratuitement.

8.6. Hashage des mots de passe

Les mots de passe des utilisateurs de la plateforme, que ce soit lecteurs ou média en ligne, doivent être hashés. La CNIL (Commission nationale de l'informatique et des libertés) recommande fortement aux plateformes telles que la nôtre de hasher le mot de passe de ses utilisateurs. De plus, il faut éviter qu'en cas de piratage de la base de données, les mots de passe ne soient utilisables par les pirates.

Pour notre prototype, nous avons choisi de prendre comme algorithme de chiffrement « Blowfish », qui fonctionne avec une clé privée, détenue par notre plateforme.

L'algorithme Blowfish fonctionne comme suit :

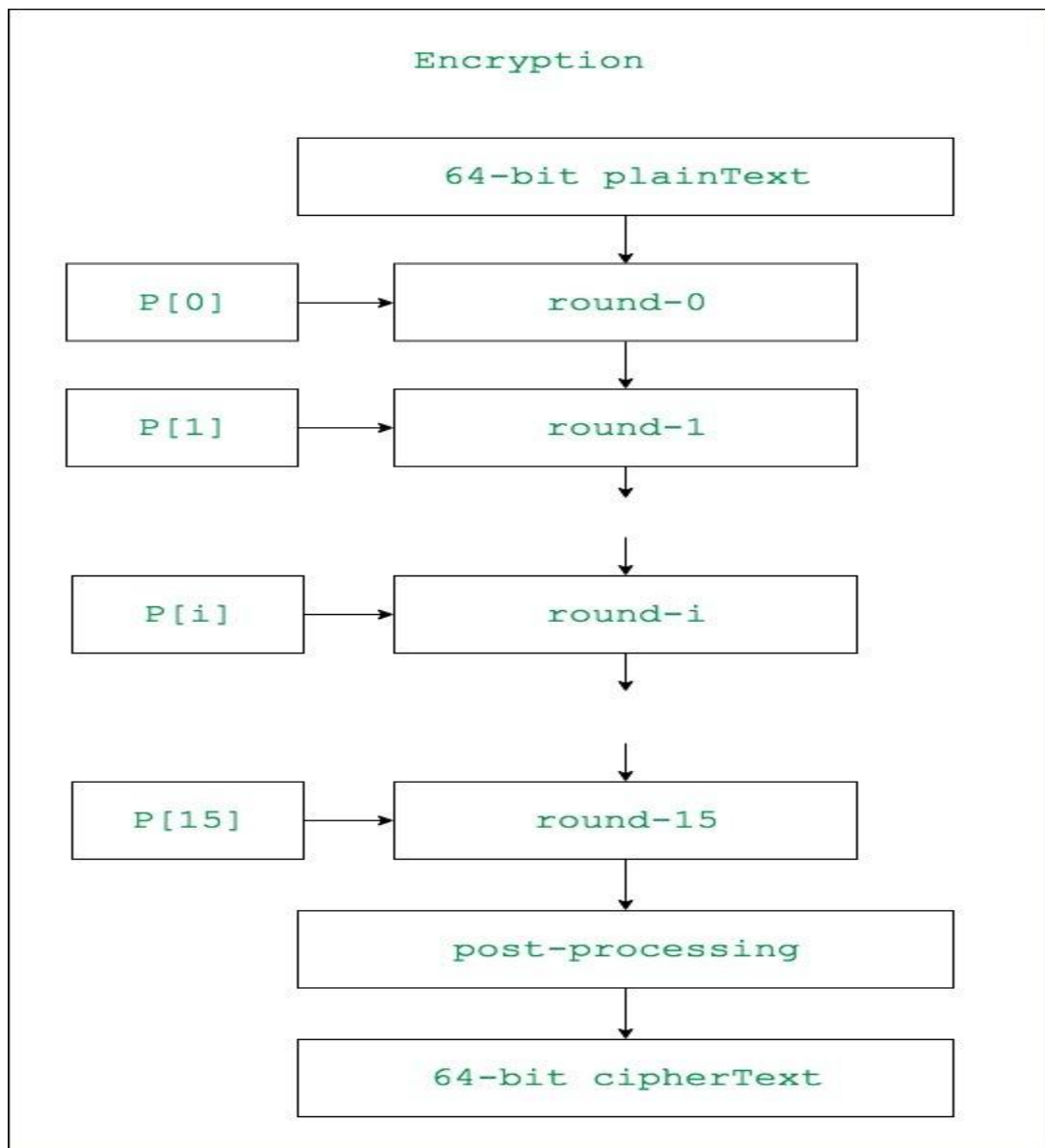


Figure 20 : Schéma algorithme Blowfish

Le hashage commence par la génération des sous-clés P de 32bits

Une suite de cryptage se fait avec, à chaque fois, la sous-clé correspondant au round (P_i), et le résultat du cryptage du round précédent.

Ensuite, des slots sont initialisés afin de garder en mémoire le résultat en reprenant par bloc les résultats des rounds. Ces slots seront ensuite assemblés suivant des opérateurs logiques ADD et XOR afin de donner ensuite un output de 32bits.

Cet output passera alors par une étape de post-processing afin d'obtenir finalement, le mot de passe hashé.

Ce mode de hashage a été choisi car il possède plusieurs avantages :

- Assez rapide après la génération des sous-clés
- Utilise un salt (Nombre aléatoire) afin de se protéger des « rainbow attack »⁵
- Génération des sous-clés assez lente, ce qui réduit l'efficacité des attaques force brute
- Répandu et libre d'utilisation

8.7. JSON Web Token

Le JSON Web Token est un moyen de transmettre une information à un autre serveur en faisant en sorte que cette information soit non modifiable par cet autre serveur, ou par un hacker qui intercepterait le token.

Encoded

PASTE A TOKEN HERE

eyJhbGciOiJIUzI1NiIsInR5cCI6IkpXVCJ9.eyJzdWIiOiIxMjM0NTY3ODkwIiwibmFtZSI6IkpvaG4gRG91IiwiaWF0IjoiYWRtaW4iOnRydWV9.TJVA95OrM7E2cBab30RMhrH DcEfxjoYZgeFONFh7HgQ

Decoded

EDIT THE PAYLOAD AND SECRET (ONLY HS256 SUPPORTED)

HEADER: ALGORITHM & TOKEN TYPE

```
{
  "alg": "HS256",
  "typ": "JWT"
}
```

PAYLOAD: DATA

```
{
  "sub": "1234567890",
  "name": "John Doe",
  "admin": true
}
```

VERIFY SIGNATURE

```
HMACSHA256(
  base64UrlEncode(header) + "." +
  base64UrlEncode(payload),
  secret
)
```

secret

☐secret base64 encoded

✔ Signature Verified

Figure 21 : Rechargement des crédits

Un JSON Web Token se découpe suivant trois chaînes de caractères consécutives, séparées par des points :

⁵ Attaque consistant à comparer des hashes de mot de passe à ceux contenus dans une table de hashes précalculée par le pirate

- La première chaîne représente un Header, du JSON encodé en base 64, dans lequel on inscrit le type de token, l'algorithme de signature utilisé, etc.
- La seconde chaîne va contenir les informations qu'on souhaite utiliser, telles que le token, les droits de la personne possédant ce token, etc. Ces informations sont formatées sous forme de « sujet » : « donnée ». Il existe des sujets particuliers, appelés claim, qui vont contenir des informations particulières, telles que la date de validité du token, le nom du serveur qui a émis le JWT, etc.
- La troisième chaîne représente la signature. Le but est de concaténer le premier et la seconde, et ensuite de les hasher avec une clé secrète. Elle permet donc de vérifier l'intégrité du token, étant donné que si la première ou la seconde chaîne est modifiée, la troisième chaîne ne correspondra plus, et on saura que le token a été changé et est donc invalide.

Il est ici utilisé pour transmettre un token d'utilisation de l'API aux différents sites de presse, et contient des informations telles que le nom du média à qui le token est destiné, la date d'expiration à partir de laquelle le token devient invalide, etc. Il est généré par l'API du projet et seule celle-ci peut décoder le contenu du token

Ainsi, chaque média partenaire possède un token différent, ce qui permet d'identifier le fournisseur d'articles de presse.

Ce token de sécurité permet d'éviter que l'API ne soit utilisée par n'importe quel serveur qui pourrait faire des requêtes afin de modifier l'état de la base de données.

9. Analytics

Une fois le prototype fonctionnel, il est intéressant d'établir quelques statistiques et d'analyser le comportement de lecteurs afin de savoir quel type d'article celui-ci préfère, pour ensuite réaliser des actions promotionnelles par exemple, notamment grâce à certaines données enregistrées dans la base de données du projet

D'autres statistiques pourraient être intéressantes pour nos médias partenaires, afin de savoir quel type d'article les lecteurs préfèrent lire, si les dossiers de presse intéressent autant que les petits articles, ou encore si durant certaines périodes de l'année, les lecteurs lisent plus d'articles en ligne.

Voici quelques requêtes qui nous ont semblé pertinentes :

- Nombre d'articles achetés par un client entre date X et date Y
→ `SELECT COUNT(*) FROM `transaction` WHERE ID_client = 2 AND `Date_achat` BETWEEN X AND Y`
- Nombre de crédits consommés par un client entre date X et date Y
→ `SELECT SUM(article.Prix) FROM transaction INNER JOIN article on transaction.ID_article = article.ID_article WHERE ID_client = 2 AND `Date_achat` BETWEEN X AND Y`
- Nombre de clients qui ont regardé un article du partenaire donné entre date X et Y
→ `SELECT COUNT(*) FROM `readarticle` WHERE `ID_article` = 1 AND `ID_provider` = 10 AND `Date` BETWEEN '2020-02-03' AND '2020-02-06'`
- Nombre de vues sur un article d'un partenaire entre date X et date Y
→ `SELECT COUNT(ID_client) FROM readarticle WHERE ID_provider = 10 AND Date BETWEEN X AND Y`

Ces requêtes permettront à MyPress de savoir quelles sont les périodes d'affluence et permettront également d'aider les sites de presse afin de savoir quels articles sont les plus souvent achetés par les lecteurs. Ils pourront alors orienter leurs rédactions vers les sujets qui intéressent le plus les utilisateurs

10. Perspective

Le prototype est fonctionnel et permet de lire des articles d'un autre site grâce au compte MyPress. Cependant, il n'est pas encore prêt à être mis en ligne directement.

Le prototype pourra être utilisé une fois intégré au média en ligne Alter Echos. Une fois un modèle économique adapté choisi, la plateforme MyPress sera en ligne et des lecteurs pourront ensuite lire des articles d'AlterEchos en utilisant leur compte MyPress. Le but ici est de montrer aux potentiels autres médias que notre projet est intéressant et également de tester la plateforme dans un véritable environnement de production.

Les objectifs et les principales étapes de ce travail de fin d'études ont été réalisés avec succès. Une étude de marché a permis d'avoir une petite idée de l'intérêt du projet et à quelle partie de la population il s'adresse.

Le site a été amélioré, autant sur la gestion des connexions que sur les modalités de paiement. La sécurité du projet est devenue une part importante de ce travail de fin d'études, et le projet est maintenant plus sûr, grâce aux protections face aux injections SQL, aux attaques XSS, etc.

Des sites de presse de test ont été créés afin de vérifier la compatibilité du projet avec un site extérieur, mais aussi pour faciliter la mise de place du projet sur la plateforme des sites de presse.

Cependant, la mise en place du projet au sein de la plateforme d'Alter Echos est pour le moment en attente, notamment à cause du confinement mis en place par l'apparition du coronavirus en Belgique. De plus, d'autres améliorations sont nécessaires avant de mettre en ligne la plateforme, certaines d'entre elles sont indiquées à la conclusion de ce travail de fin d'études

11. Conclusion

Ce travail de fin d'études avait pour ambition de présenter un prototype de plateforme en ligne capable de servir d'intermédiaire entre un lecteur et un fournisseur d'article de presse afin de centraliser l'accès à la presse en ligne. Le but était de montrer la faisabilité du projet, et d'aborder les aspects de sécurité afin de protéger les lecteurs, mais aussi les fournisseurs partenaires.

Il a fallu d'abord modéliser le problème afin de proposer une architecture viable qui puisse gérer les connexions des clients et partager des informations aux fournisseurs d'articles de presse afin de gérer les accès. Ensuite, j'ai réalisé un prototype répondant aux attentes de l'entreprise, et débouchant sur une implémentation fonctionnelle de l'architecture proposée. Enfin, j'ai testé le prototype en l'interconnectant avec plusieurs sites de test ce qui m'a permis de réparer les problèmes qu'ils ont mis en lumière.

Le prototype réalisé ici est encore à améliorer, autant pour augmenter la sécurité de celui-ci que pour simplifier son utilisation. De nombreuses nouvelles fonctionnalités pourront aussi être implémentées :

- La partie Front-end a été mise de côté, il faut donc encore la développer afin de rendre le site attrayant et plus simple pour l'utilisateur
- L'ID des clients est enregistré comme un simple entier. Il faut hasher l'ID des clients, afin qu'un pirate ne puisse se connecter sur le compte d'un autre en modifiant un de ses cookies.
- Une page de contact permet aux médias intéressés par notre plateforme de contacter MyPress afin de mettre en place une collaboration. Mais il faut configurer un serveur mail afin d'envoyer directement un mail à MyPress lorsqu'un média est intéressé.
- Lorsque le token d'un média expire, celui-ci doit alors se ré-authentifier auprès de l'API afin de recevoir un nouveau token. Il faudrait rajouter une fonction qui permettrait à l'API de donner automatiquement un nouveau token à un média lorsque celui-ci est proche de sa date d'expiration.

- L'ajout d'un article dans la base de données se fait à la main, il faudrait permettre aux médias de faire une requête API afin de stocker automatiquement un nouvel article.

Il serait également possible de changer le système des articles et d'utiliser une signature numérique : Le média encode les informations sur son article (avec JWT par exemple), ensuite il l'inscrit dans son bouton JavaScript, et lorsque le lecteur clique sur le bouton, MyPress procède à l'achat de l'article avec les données du JWT

Ainsi, il ne sera plus nécessaire de stocker chaque article du média, et un pirate ne peut pas modifier les informations protégées par la JWT

- Il faut ajouter un certificat HTTPS à l'API du projet, pour le moment, seul le site de la plateforme est en HTTPS
- Prendre en compte la convention GDPR, modifier la base de données en conséquence et ajouter des boutons pour modifier ou supprimer les informations personnelles des utilisateurs

Ce travail de fin d'études m'a appris à concevoir une plateforme en ligne, en partant de la conception de l'architecture, jusqu'à l'intégration d'un prototype dans des conditions réelles, en respectant un cahier des charges en plus des demandes de l'entreprise.

Il m'a également poussé à créer un code flexible, afin de respecter les demandes de l'entreprise, qui fluctue en fonction du temps et des opportunités du projet. Il m'a montré l'importance d'une bonne méthode de travail et de réunions régulières, afin de vérifier aussi souvent que possible l'état du projet et ainsi savoir ce qu'il faut modifier et quels sont les prochains objectifs courts termes.

Ce projet de centralisation de la presse arrive dans un monde où l'accès à l'information se fait en grande partie sur Internet, où les prochaines générations se renseignent le plus souvent à travers diverses sources. Le modèle de l'achat à la pièce semble alors intéressant. Plusieurs médias ont changé et se sont adaptés à la presse en ligne, mais très peu ont changé leur modèle économique vers un achat à la pièce.

L'adoption de notre projet par Alter Echos pourrait alors convaincre les autres médias de l'intérêt de MyPress, et d'un système économique d'achat d'article à la pièce.

12. Bibliographie

- [1] Pixis, Hackndo, *L'attaque XSS*, Septembre 2017, <https://beta.hackndo.com/attaque-xss/>
- [2] Wikipedia, *HTTP cookie*, Mai 2011, https://en.wikipedia.org/wiki/HTTP_cookie
- [3] Openclassroom, *Protégez-vous efficacement contre les failles web*, Mars 2020 <https://openclassrooms.com/fr/courses/2091901-protegez-vous-efficacement-contre-les-failles-web/2680180-linjection-sql>
- [4] Aliou Saw, AfricanFreelancers, *L'architecture MVC*, Octobre 2018 <http://africanfreelancers.net/blog/larchitecture-mvc>
- [5] Jwt, *Introduction to JSON Web Tokens*, <https://jwt.io/introduction/>
- [6] ParthDutt, GeeksforGeeks, *Understanding Rainbow Table Attack*, <https://www.geeksforgeeks.org/understanding-rainbow-table-attack/>
- [7] UKEssays. *Blowfish Algorithm Advantages and Disadvantages*, November 2018, <https://www.ukessays.com/essays/computer-science/blowfish-algorithm-history-and-background-computer-science-essay.php?vref=1>
- [8] Howard Poston, Keshav Dhandhanian, Commonlounge, *Blowfish: The first well-known encryption algorithm in public domain*, <https://www.commonlounge.com/discussion/d95616beec148daaa23f35178691c35>
- [9] Certbot, <https://certbot.eff.org/lets-encrypt/ubuntuxenial-apache>
- [10] TutorialRepublic, <https://www.tutorialrepublic.com/php-tutorial/php-mysql-select-query.php>
- [11] Stripe, <https://stripe.com/docs/development>
- [12] Developpez, *La gestion des cookies en JavaScript*, <https://ppk.developpez.com/tutoriels/javascript/gestion-cookies-javascript/>
- [13] Bootstrap, *Introduction*, <https://getbootstrap.com/docs/4.5/getting-started/introduction/>
- [14] Eric, PresseCitron, *Presse en ligne : pourquoi payer un abonnement quand on veut juste lire un article ?*, 19 aout 2019, <https://www.presse-citron.net/presse-en-ligne-pourquoi-payer-un-abonnement-quand-on-veut-juste-lire-un-article-le-kick-off/>

- [15] RTBF *Les jeunes s'informent d'abord sur les réseaux sociaux (étude)*, 03 août 2018
https://www.rtb.be/tendance/bien-etre/psycho/detail_les-jeunes-s-informent-d-abord-sur-les-reseaux-sociaux-etude?id=9982804
- [16] AbhayBhat , GeeksforGeeks, *Blowfish Algorithm with Examples*,
<https://www.geeksforgeeks.org/blowfish-algorithm-with-examples/>
- [17] Kassandra Perch, SitePoint, *Securing Your IoT Devices and Services with JSON Web Tokens*, 05 Juillet 2016
<https://www.sitepoint.com/securing-your-iot-devices-and-services-with-json-web-tokens>
- [18] Marsner, *Why Test-Driven Development (TDD)*,
<https://marsner.com/blog/why-test-driven-development-tdd/>

13. Annexes

13.1. Accueil et ses boutons

La page d'accueil du site du prototype se présente comme suit :



Il possède plusieurs boutons :

- MyPress : Permet de revenir à l'accueil
- Formules : Affiche les différentes formules de rechargement de crédits



- Partenaires : Affiche les médias partenaires du projet



- Contact : Permet au site de presse de prendre contact avec le groupe MyPress

MyPress
FORMULES
PARTENAIRES
CONTACT

CONNEXION

Nom :

Courriel :

Objet :

Message :

ENVOYER

- Connexion : Permet aux utilisateurs de se connecter à leur compte MyPress

Connexion



LecteurMedia

Nom de Compte :

Mot de Passe :

CONNEXION

PAS ENCORE INSCRIT ?

13.2. Inscription

Pour s'inscrire, le lecteur doit entrer ses informations de connexions et son adresse

MyPress

FORMULES

PARTENAIRES

CONTACT

CONNEXION

S'inscrire en tant que lecteur

E-mail :

Mot de Passe :

Retaper le mot de passe :

Type de paiement :

Adresse :

INSCRIPTION

13.3. Lecteur connecté

Une fois connecté, le client a accès à ses informations, et un bouton lui permet de changer son mot de passe

MyPress

FORMULES

PARTENAIRES

CONTACT

DÉCONNEXION

Profil

Nom de compte : zbeb;

Crédits : 201

Paiement : B

Adresse : adroouuse

CHANGER PASSWORD

RECHARGER

Portefeuille

Transactions :

ID : 54 Article : 5 Media : Leu_Sware Date : 2019-12-05

Factures :

ID : 7 Objet : Pack 25 articles Montant : 20.00 Date : 2020-01-28 18:53:39

ID : 8 Objet : Pack 25 articles Montant : 20.00 Date : 2020-01-28 18:54:29

ID : 9 Objet : Pack 25 articles Montant : 20.00 Date : 2020-01-28 19:38:30

ID : 10 Objet : Pack 25 articles Montant : 20.00 Date : 2020-01-28 19:39:57

ID : 11 Objet : Pack 10 articles Montant : 10.00 Date : 2020-01-28 19:49:25

Un bouton lui permet de se déconnecter et est présent sur toutes les pages du site du projet.

Pour recharger son compte, un lecteur choisi une des offres de rechargement

MyPress

FORMULES PARTENAIRES CONTACT

DÉCONNEXION

Rechargement des crédits

Veuillez choisir le montant de crédits :

☐ 1 article : 1e

☐ Pack 10 articles : 10e (Populaire)

☐ Pack 25 articles : 20e

VALIDER

Une fois choisi, il doit entrer ses coordonnées bancaires, qui seront envoyés à la plateforme Stripe qui se chargera de faire le virement

MyPress

FORMULES PARTENAIRES CONTACT

DÉCONNEXION

Charge €20 with Stripe

Item Name: Pack 25 articles

Price:20 EUR

NAME

EMAIL

CARD NUMBER

EXPIRY DATE

CVC CODE

MM

YYYY

SUBMIT PAYMENT

Si le paiement est réussi, une page s'affiche

Your Payment has been Successful!

Payment Information

Reference Number: 12

Paid Amount: 20 eur

Payment Status: succeeded

Product Information

Name: Pack 25 articles

Price: 20 EUR

[Back to home Page](#)