

LE MANS UNIVERSITÉ

MASTER INFORMATIQUE 1ÈRE ANNÉE

MODULE 178UD07 - CONCEPTION ET
DÉVELOPPEMENT DE LOGICIEL

Rapport projet HRhelper

Baptiste DUBIN
Anaïs MOTTIER
Simon PLANCHENAULT
Julien PROUDY

Enseignant :
Claudine PIAU-TOFFOLON
Tuteur :
Madeth MAY

[Dépôt GitHub](#)

Année universitaire 2021/2022

Table des matières

1	Cadre et but du document	2
2	Le projet	2
3	Analyse des besoins	3
3.1	Personas	3
3.1.1	Pauline SEVIER, chargé de recrutement	3
3.1.2	Jean PEZARD, responsable des ressources humaines	3
3.2	Cas d'utilisation	4
3.3	Scénarios d'utilisation	4
3.4	Users stories	6
4	Le scraping	6
4.1	Qu'est ce que c'est ?	6
4.2	Problématique	7
5	Notre approche du projet	8
5.1	Les données	8
5.2	L'application Web	9
5.3	Éléments de conception et de programmation	10
5.3.1	Scraping	11
5.3.2	Stockage	11
5.3.3	Les indicateurs	12
5.3.4	Angular	12
6	Gestion de projet	13
6.1	Organisation du groupe	13
6.2	Outils utilisés	14
7	Réalisation	14
8	Conclusion	15
9	Annexes	16
9.1	Traitement des données issus du scraping du site Pôle Emploi	16
9.2	Fiche d'activité DUBIN Baptiste	18
9.3	Fiche d'activité MOTTIER Anaïs	19
9.4	Fiche d'activité PLANCHENAULT Simon	20
9.5	Fiche d'activité PROUDY Julien	21
	Lexique	22

1 Cadre et but du document

Dans le cadre de notre cursus universitaire en Master informatique, à Le Mans Université, dans le module 178UD07 : Conception et Développement de Logiciel, nous avons réalisé un projet informatique pour un client, Madeth MAY.

Ce document a pour but de retracer la réalisation du projet d'un point de vue technique mais également d'un point de vue gestion de projet.

2 Le projet

Le monde du travail est un monde en constante évolution, des milliers de postes sont à pourvoir tous les jours. C'est un marché très actif, il faut savoir être rapide et réactif dans ce milieu.

Les chargés de recrutement et responsables des ressources humaines (RH) passent énormément de temps à chercher des candidats pour différents postes avec des profils plus ou moins définis. Ces recherches peuvent être longues et coûteuses pour une entreprise, en effet plus le recruteur passe de temps sur la recherche d'un candidat plus l'entreprise perd de l'argent. Les entreprises sont donc à la recherche de solutions pouvant permettre à leurs recruteurs d'être plus efficaces et rapides dans leurs recherches.

Comment permettre aux recruteurs d'être plus efficaces dans la recherche de profils ?

Comment trouver des candidats qui correspondent réellement aux critères pour un poste ?

Le projet HRhelper est un projet informatique qui cherche à répondre à ces problématiques. Il est destiné à faciliter la vie des recruteurs et responsables des ressources humaines qui veulent trouver des profils adéquats aux postes qu'ils proposent en se basant sur des filtres, mots clés et indicateurs (graphiques) depuis une page web.

3 Analyse des besoins

Afin de réaliser au mieux le projet HRhelper l'équipe a fait une analyse des besoins en définissant des personas, les scénarios d'utilisation, les cas d'utilisation et Users Stories en collaboration avec le client.

3.1 Personas

3.1.1 Pauline SEVIER, chargé de recrutement

Pauline SEVIER est une chargée de recrutement de 29 ans travaillant chez MMA, un groupe d'assurance mutuelle française. Elle est très familière avec les nouvelles technologies, en effet avant de faire sa licence en ressources humaines (bac +3), elle a fait un parcours qui n'a pas abouti en informatique.

A ce jour, elle souhaite avoir un outil qui lui permet, à partir d'une base de données de CV sur des sites de recrutements comme LinkedIn ou encore Indeed et Pôle emploi, de récupérer des profils et pouvoir, à partir de filtres qu'elle peut définir, trouver les meilleurs profils en fonction de sa recherche. Grâce à cet outil, elle sera capable de trouver les meilleurs candidats pour un poste et ainsi recruter uniquement de bons éléments pour son entreprise. Mais, consciente de l'importance de la vie privée sur internet, elle souhaite que cet outil respecte les règles éthiques fondamentales.



3.1.2 Jean PEZARD, responsable des ressources humaines



Jean PEZARD est responsable des ressources humaines dans une société d'électricité. Il a 53 ans et aime être en relation directe avec les employés de l'entreprise dans laquelle il travaille. Le secteur du bâtiment ne connaît pas la crise et son entreprise est en première ligne. Cette dernière est en plein essor et à besoin de beaucoup de main d'œuvre qualifiée supplémentaire. Cependant, il a peu de temps à consacrer à la recherche de candidats, il lui faut donc être le plus efficace possible.

Jean a déjà essayé d'utiliser des sites internet afin de l'aider dans ses recherches mais malheureusement pour lui ces sites sont souvent compliqués à prendre en main pour les personnes qui ne sont pas à l'aise avec l'informatique, comme lui. Pourtant, utiliser une application web simple et rapide, qui trouve des candidats, lui serait d'une très grande aide.

3.2 Cas d'utilisation

L'équipe a spécifié différents cas d'utilisation de la page web pour un utilisateur.

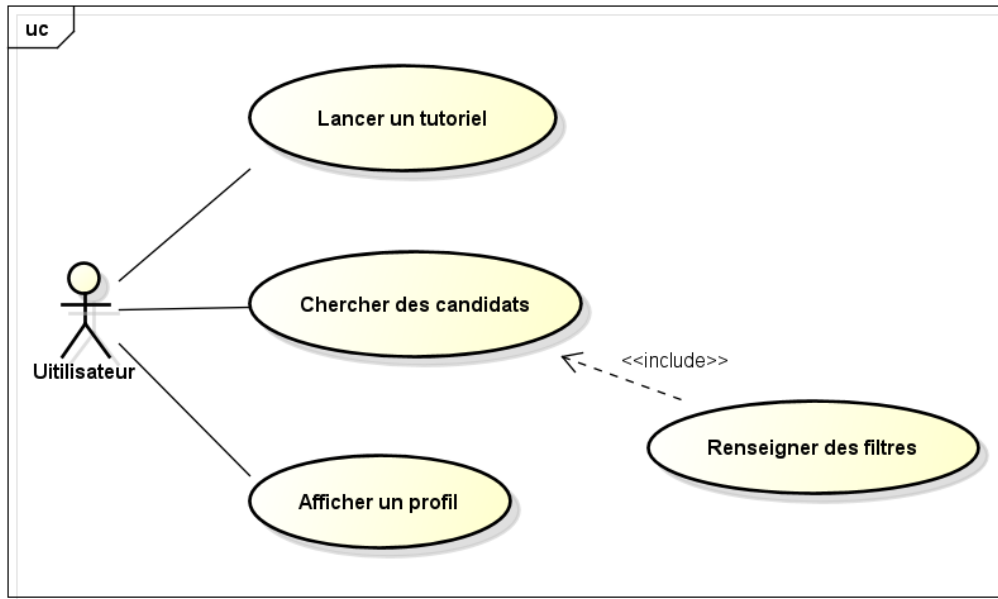


FIGURE 1 – Diagramme de cas d'utilisation HRhelper

L'application web HRhelper permettra à l'utilisateur de :

- **lancer un tutoriel** afin de comprendre le fonctionnement du site et comment l'utiliser,
- **chercher des candidats** : pour cela il faudra que l'utilisateur renseigne des filtres tels que le niveau d'étude, la disponibilité, le permis, etc. Il pourra également utiliser des mots clefs,
- mais aussi **afficher le profil** détaillé d'un candidat : expériences, formations, compétences...

3.3 Scénarios d'utilisation

Scénario n°1 :

Un recruteur est à la recherche d'un(e) conseiller(e) clientèle pour une grosse chaîne de distribution alimentaire à domicile. Il recherche une personne ayant fait des études de commerce, possédant le permis de conduire, pour qu'il puisse démarcher les clients, mais également un bon niveau en anglais et un bon relationnel.

Pour trouver un candidat correspondant à cette recherche le recruteur va se rendre sur l'application HRhelper puis suivre les étapes suivantes :

- cliquer sur le bouton "COMMENCER" ou, s'il découvre le site, il peut lancer le tutoriel,
- cliquer sur le bouton filtre (en bas à droite) afin de créer son filtrage,
- activer la catégorie "Permis",
- sélectionner "Anglais" dans la partie "Langues",
- inscrire "relationnel", "commerce" et "conseiller clientèle" dans le champs "Mot-clé",
- et enfin cliquer sur le bouton "Valider",

Une fois ceci fait, la liste des profils apparaît, il peut maintenant la parcourir et afficher les différents profils.

Scénario n°2 :

Une responsable RH recherche un(e) homme/femme de ménage en urgence pour une société d'entretien de locaux professionnels. Elle souhaite un(e) candidat(e) disponible immédiatement, dynamique..

Pour trouver un candidat correspondant à cette recherche la responsable RH va se rendre sur l'application HRhelper puis suivre les étapes suivantes :

- cliquer sur le bouton "COMMENCER", ou si elle découvre le site, elle peut lancer le tutoriel,
- cliquer sur le bouton filtre (en bas à droite) afin de créer son filtrage,
- cocher "immédiate" dans la partie "Disponibilité",
- inscrire "dynamique" et "ménage" dans le champs "Mot-clé",
- et enfin cliquer sur le bouton "Valider",

Une fois ceci fait, la liste des profils apparaît, elle peut maintenant la parcourir et afficher les différents profils.

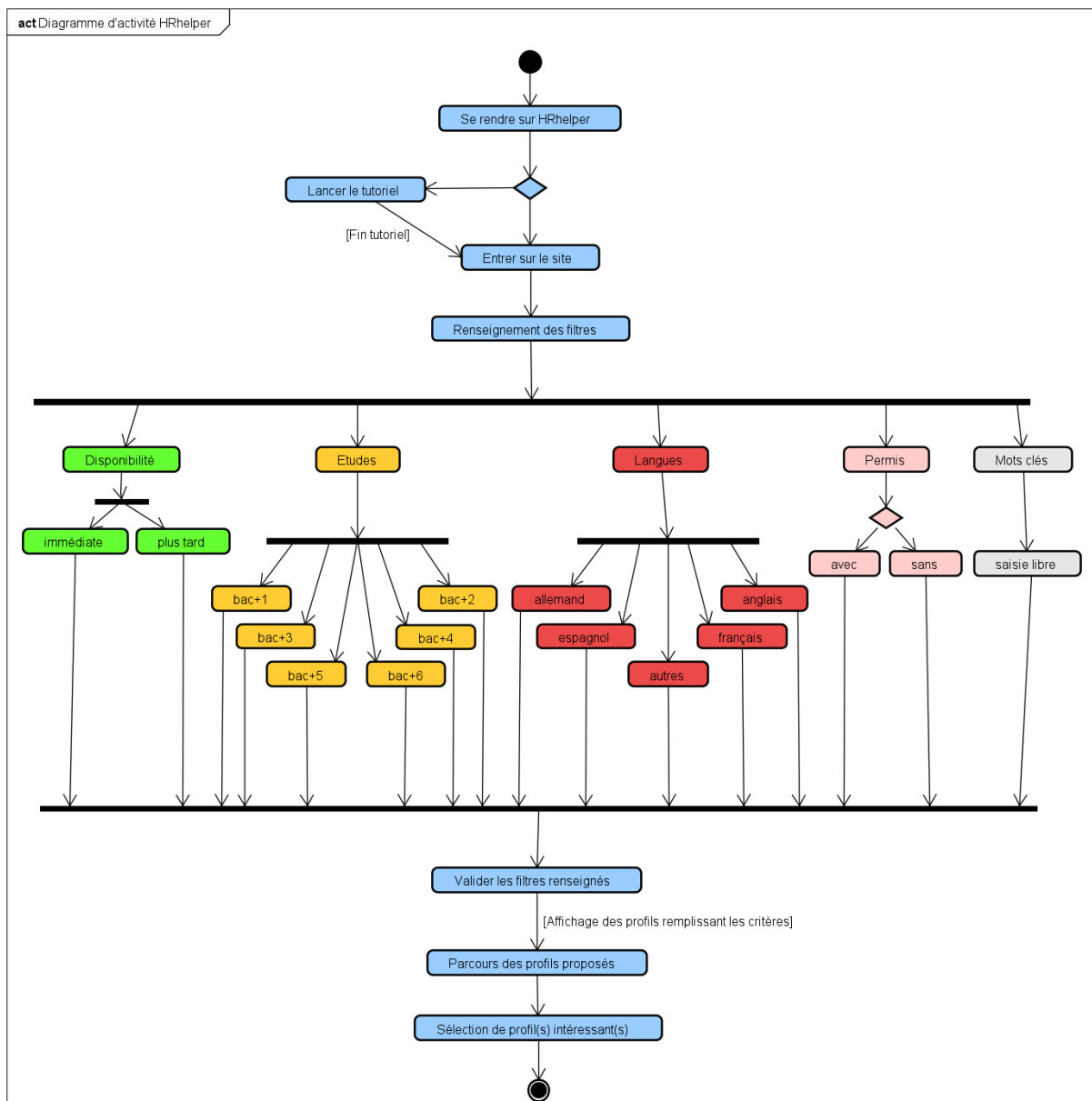


FIGURE 2 – Diagramme d'activité HRhelper

3.4 Users stories

Vous trouverez ci-dessous les différentes Users Stories de l'application.

- **User Story 1 : Trouver des candidats.**
En tant qu'utilisateur je souhaite pouvoir trouver des candidats à la recherche d'un emploi.
- **User Story 2 : Sélectionner des candidats par des filtres.**
En tant qu'utilisateur je souhaite pouvoir filtrer les candidats via des filtres et mots clefs afin d'avoir uniquement des candidats qui correspondent à ma recherche.
- **User Story 3 : Afficher un profil de candidat.**
En tant qu'utilisateur je souhaite pouvoir afficher le profil d'un candidat (compétences, expériences, formations, langues, etc).
- **User Story 4 : Avoir des indicateurs sur les candidats.**
En tant qu'utilisateur je souhaite avoir un graphique représentant le candidat pour pouvoir me faire un avis rapide et synthétique sur celui-ci et savoir s'il convient au poste.
- **User Story 5 : Avoir accès à un tutoriel.**
En tant qu'utilisateur je souhaite pouvoir avoir accès à un tutoriel afin de savoir comment fonctionne l'application.
- **User Story 6 : Avoir accès aux informations des candidats de façon éthique.**
En tant qu'utilisateur je souhaite pouvoir avoir accès aux informations des candidats de manière éthique : pas de discrimination possible et d'informations personnelles à son sujet afin d'avoir un avis impartial et neutre, uniquement basé sur les compétences et qualités de la personne.

4 Le scraping

4.1 Qu'est ce que c'est ?

Le scraping ou data scraping, qu'on traduit en français par extraction de contenus, est une technique permettant d'extraire de l'information depuis une source telle que des sites internet, des documents... et ce de manière automatique. L'extraction des ces informations se fait par des scripts et programmes informatiques.

Le data scraping peut être utilisé pour différentes raisons telles que rechercher des contenus Web, comparer des tarifs (trouver le meilleur prix, rester compétitif par rapport aux concurrents...), mais aussi mener une étude de marché, chercher des canaux de vente, ou encore transférer des données entre plusieurs sites et référencer des sites web.

Il existe plusieurs types de scraping :

- le web scraping se fait depuis une page web,
- le report mining quand à lui se fait depuis un fichier texte,
- et enfin le screen scraping se fait à partir d'un écran.

Le web scraping est la méthode la plus courante, il permet d'extraire des données depuis des pages web grâce à un navigateur internet ou protocole HTTP.

Le web scraping est réalisé en trois étapes :

1. **La requête** (le fetching) : récupération de la page web par téléchargement.
2. **Le parsing** : traitement du code HTML de la page internet afin d'extraire uniquement ce dont on a besoin (lecture, récupération, structuration).
3. **Le stockage** : stockage des informations au format choisi, exemple dans une base de données, ou dans un fichier au format JSON, CSV, XML... afin de pouvoir par la suite travailler dessus et les réutiliser.

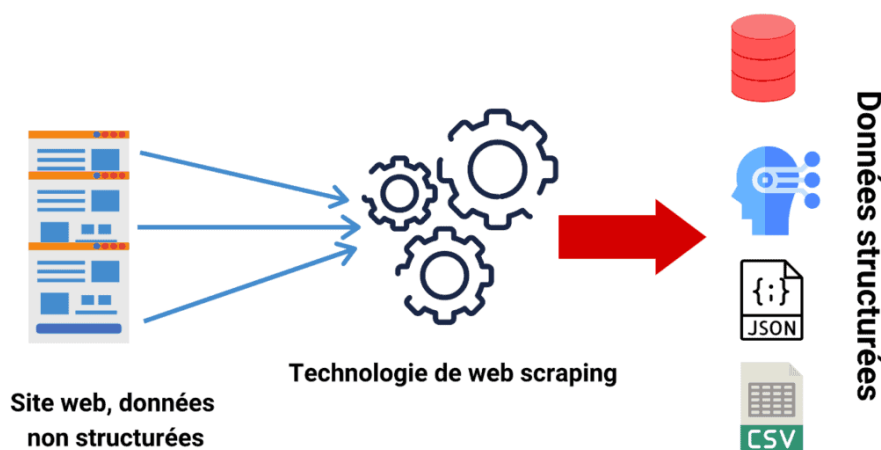


FIGURE 3 – Schéma scraping

Cette technique permet de réutiliser du contenu, c'est pourquoi certains l'associent à du pillage, copiage et pompage de données. C'est la raison pour laquelle de plus en plus de sites et moteurs de recherche détectent la technique du scraping et la condamnent.

4.2 Problématique

Comme introduit précédemment, la technologie du scraping est très controversée. L'auteur du contenu scrapé n'a aucun moyen de connaître le contexte dans lequel sera utilisé son contenu, de manière bienveillante ou malveillante, ce qui pose un problème d'un point de vue éthique.

Quels types de données sont scrapées ? Comment vont-elles être utilisées ? Et dans quel but sont-elles récupérées ? Vont-elles être détournées ? Où sont stockées ces informations ? Qui en aura l'accès ?

Lorsqu'un robot/programme scrap un site il est impossible pour ce dernier de connaître les réponses aux questions précédentes, c'est pourquoi les moteurs de recherche et sites internet sont de plus en plus méfiants et cherchent continuellement à améliorer leur détection du scraping. Pour cela, certains mettent en place des bloqueurs de scrapers (CAPTCHAs, page Honeypot...), ils indiquent également dans leurs Conditions Générales d'Utilisations (CGU) qu'il est interdit d'utiliser des technologies de ce type sur leur site. Avec ces restrictions il est devenu difficile de scraper des pages web sans se faire bloquer/bannir ou avoir des requêtes nulles.

5 Notre approche du projet

Afin de pouvoir répondre à cette demande et aider la cible utilisateur dans sa recherche de candidats. L'équipe a imaginé une application web, nommé HRhelper. Cette application a pour but d'éviter aux recruteurs et responsables RH de parcourir tous les sites ayant des CVthèque un à un. Avec HRhelper, ils auront accès à l'ensemble des candidats sur une même application, une manière plus simple, rapide et efficace de trouver le meilleur élément pour le poste. HRhelper permet de filtrer et afficher des profils de candidats en fonction de filtres renseignés par l'utilisateur. Ce dernier peut visualiser chaque profil sous la forme d'un graphique (radar chart) afin de pouvoir se faire un premier avis assez rapidement sur le candidat sélectionné et donc faciliter la prise de décision.

Afin de réaliser ce projet, plusieurs étapes sont importantes :

1. Concevoir et développer l'algorithme de scraping des données depuis les pages web (LinkedIn et Pôle Emploi) sur lesquels se trouve les profils de candidats.
2. Transformer les données (profils) récupérées dans un format exploitable par un système informatique et les stocker.
3. Traiter ces données.
4. Développer et concevoir l'application web.

5.1 Les données

La récupération des données et leurs traitements ont été mûrement réfléchis au sein de l'équipe. En effet ce type de données (informations personnelles) est très délicat à manipuler et l'utilisation de la technique du scraping accentue cet aspect.

Il a été décidé de scraper les sites LinkedIn et Pôle emploi qui constituent à eux deux une grosse base de données de candidats. Le projet HRhelper se veut le plus éthique possible, en accord avec cette valeur, uniquement les informations dites "non sensibles" et ne pouvant pas donner d'indications sur l'identité du candidat, afin de garder l'anonymisation et éviter tout type de discrimination, seront scrapées et utilisées. Par conséquence, les informations suivantes ont été sélectionnées pour constituer un profil :

- expériences
- diplômes/formations
- compétences
- langues
- permis
- disponibilité
- date de mise à jour du profil

Les informations de types : nom, prénom, âge, nationalité, photo, sexe, loisirs, contacts, etc, ne sont pas affichées.

Afin de sécuriser les données, elles sont stockées dans une base de données (BDD) après avoir été récupérées dans un JSON à la suite d'un scraping.

5.2 L'application Web

HRhelper utilise une interface web qui se veut la plus interactive, intuitive et responsive possible. Elle sera composée de plusieurs parties :

- **L'accueil** : composé d'un tutoriel afin d'expliquer l'utilisation du site.
- **La liste des candidats** : la liste des candidats s'affiche une fois les filtres renseignés.
- **Les filtres** : le renseignement de ceux-ci est obligatoire afin de sélectionner uniquement les candidats correspondant à la recherche de l'utilisateur. Les filtres possibles sont :
 - la disponibilité : *immédiate ou plus tard*,
 - les études : *bac+1, bac+2, bac+3, bac+4, bac+5, bac+6 et/ou autres*,
 - les langues : *allemand, anglais, espagnol, français et/ou autres*,
 - le permis : *avec ou sans*,
 - des mots clés : *champs de saisie libre*.
- **Le candidat** : affichage du profil du candidat avec les informations présentées au dessus dans la section 5.1 et une synthèse de celui-ci grâce à un graphique de type radar chart.

Des maquettes de modélisation du site web ont été réalisées afin d'aider à la réalisation de l'application.



FIGURE 4 – Maquette accueil HRhelper

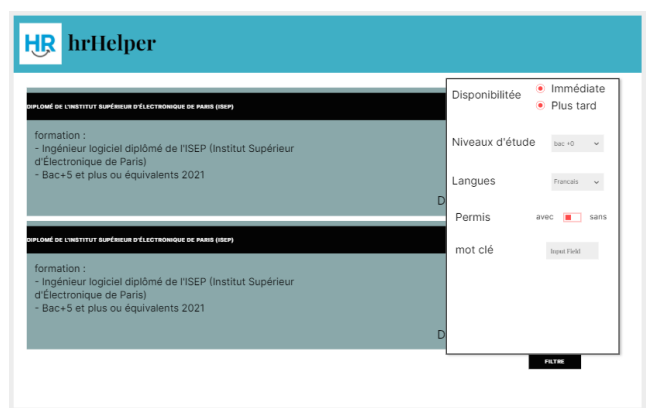


FIGURE 5 – Maquette filtres HRhelper



FIGURE 6 – Maquette liste HRhelper

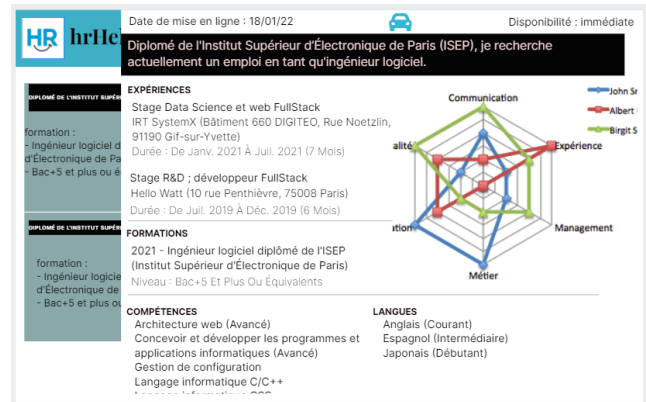


FIGURE 7 – Maquette profil candidat HRhelper

5.3 Éléments de conception et de programmation

Pour avoir une vue d'ensemble de la conception du projet une architecture logicielle a été conçue.

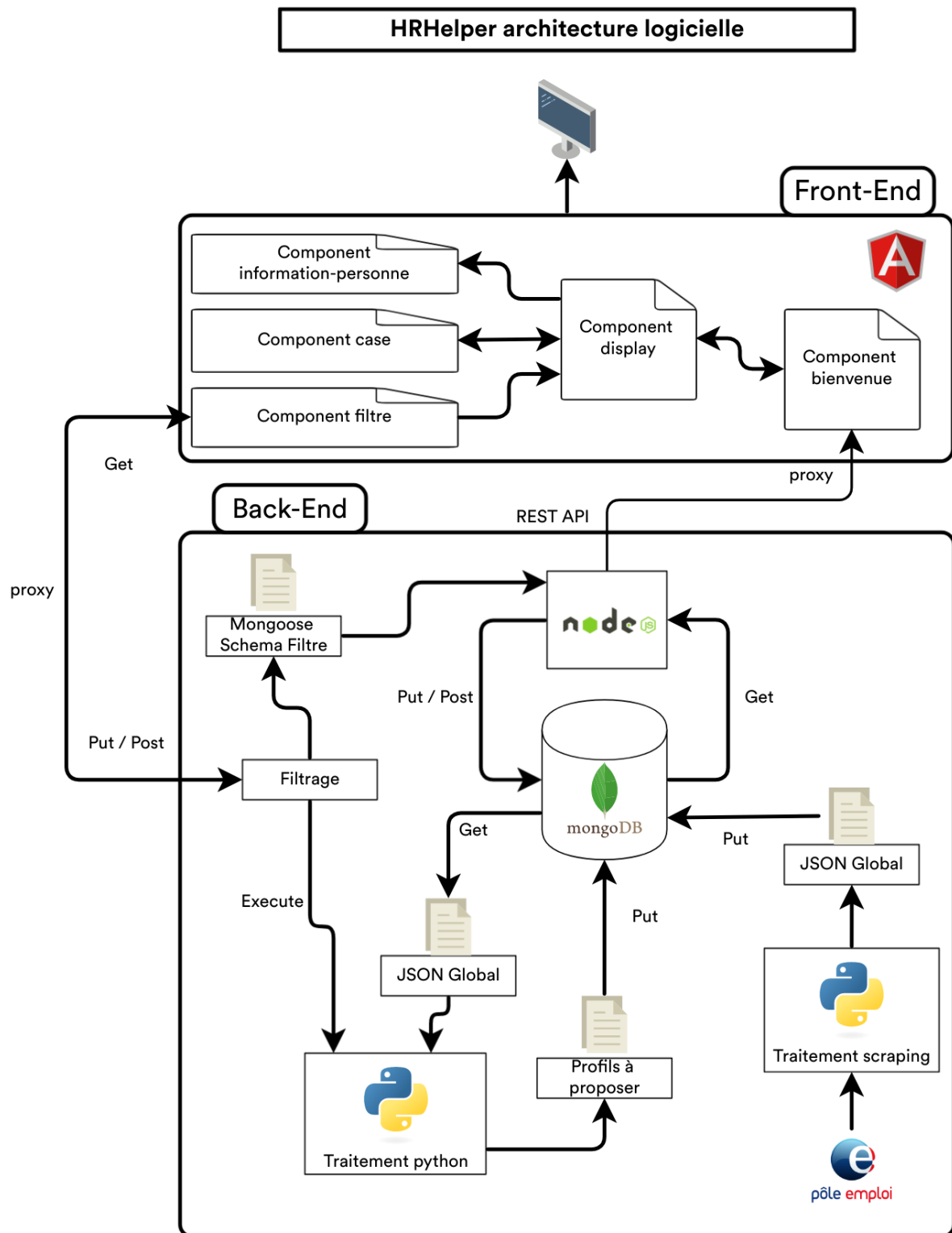


FIGURE 8 – Schéma architecture logicielle

5.3.1 Scraping

Afin de scraper les deux CVthèques, deux frameworks sont utilisés : Scrapy pour LinkedIn et Selenium pour Pôle Emploi.

Nous avons utilisé Scrapy pour scraper les données de LinkedIn puisqu'il permet une récupération rapide d'une grande quantité de données. Au début du projet, nous pensions qu'il était possible de scraper tous les profils de LinkedIn, mais, cette plateforme permet uniquement d'avoir accès au profils des personnes de son réseau ou des personnes autorisant la consultation de leurs profils (profil public).

Nous avons utilisé Selenium pour scraper les données de Pôle Emploi puisque Scrapy permet difficilement de naviguer sur les sites qui utilisent JavaScript. Et ce malgré que Selenium ait un temps de récupération des profils beaucoup plus long que Scrapy.

Pour des raisons d'efficacité seulement 25 profils sont scrapés à la fois. A la fin du scraping les informations se trouvent dans un fichier JSON.

5.3.2 Stockage

Les données récupérées suite au scraping, dans le fichier JSON sont par la suite stockées dans une base de données (BDD) créée sur MongoDB. La BDD, appelée profilesDB, utilisée par l'application contient 3 tables : profiles, filtres et retour_filtres. Chacune à une utilisation bien précise :

- la **table profiles** qui contient tous les profils scrapés,
- la **table filtre** elle, contient le filtre renseigné par l'utilisateur,
- et enfin la **table retour_filtres** qui est composée des profils correspondant à la recherche de l'utilisateur.

Collection Name	Documents	Documents Size	Documents Avg	Indexes	Index Size	Index Avg
filtres	1	129B	129B	1	36KB	36KB
profiles	14	21.97KB	1.57KB	1	20KB	20KB
retour_filtres	3	1.92KB	657B	1	20KB	20KB

FIGURE 9 – Base de données profilesDB

Pour le moment la BDD n'est pas mise à jour, pour cela il faudrait refaire un scraping à chaque nouvelle recherche ce qui serait très coûteux en temps.

5.3.3 Les indicateurs

Pour permettre à l'utilisateur d'avoir une vue globale d'un profil de candidat, un graphique de type radar chart sont développés pour chacun des profils correspondant à la recherche. Cette fonctionnalité permet à l'utilisateur d'être plus efficace et rapide dans ses recherches. Le graphique prendra en compte 3 indicateurs :

- le niveau en langue vivante,
- les expériences professionnels,
- et les formations.

Pour avoir plus de détail sur la manière dont sera conçu les différents indicateurs du radar chart référez-vous à l'annexe 1 : "Traitement des données issus du scraping du site Pôle Emploi".

5.3.4 Angular

L'application web utilise la technologie d'Angular reconnue pour ses nombreux avantages : facile à apprendre, utiliser et tester, le code est réutilisable, la liaison de données bidirectionnelles... La page web se compose de plusieurs composants, vous trouverez ci-dessous un schéma récapitulatif de l'organisation des composants et les différentes données qu'ils se transmettent entre eux.

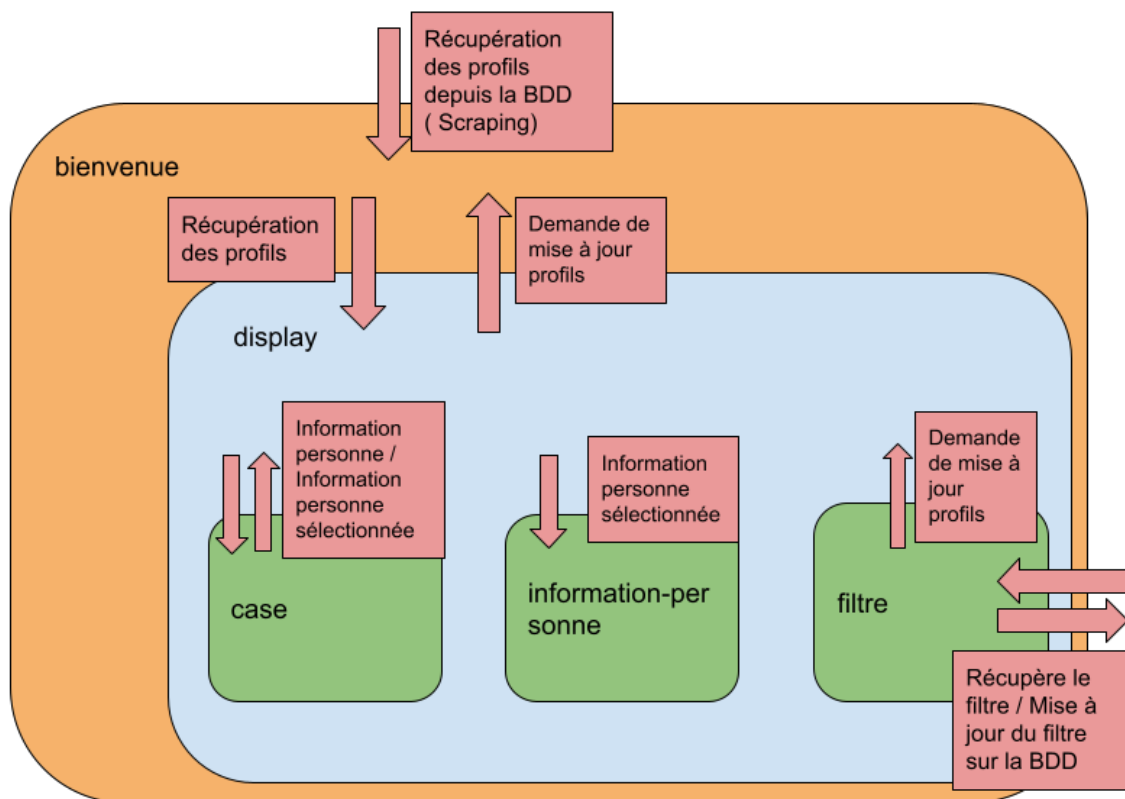


FIGURE 10 – Schéma architecture Angular

Angular est relié à la BDD par un serveur : Node JS API REST pour récupérer et également envoyer des données suite à une demande de filtrage des candidats par l'utilisateur via le composant filtre.

6 Gestion de projet

La réalisation d'un projet nécessite une organisation et une gestion. Pour HRhelper un tableau de type [Gantt](#) a été créé afin d'avoir un macro-planning. L'équipe a souhaité suivre une méthode de génie logiciel, la méthode SCRUM. Pour cela, il a fallu découper le projet en plusieurs sprints composés chacun de Backlogs déterminés auparavant. Lors de chaque réunion de travail, un daily meeting était fait afin de savoir où chacun en était, si des difficultés étaient rencontrées, si le projet nécessitait des réajustements, etc. Un référent projet a été désigné afin de simplifier les échanges avec le client et le responsable du module universitaire : Anaïs MOTTIER.

6.1 Organisation du groupe

Chaque membre du groupe a travaillé sur des tâches différentes comme vous pouvez le voir dans le tableau ci-dessous.

Membre équipe	Tâches
DUBIN Baptiste	<ul style="list-style-type: none">• Génie Logiciel : backlog, diagramme de classe (version 1)• Scraping LinkedIn et Pôle emploi• Codage de l'indicateur (radar chart)
MOTTIER Anaïs	<ul style="list-style-type: none">• Génie Logiciel : backlog, sprint, Gantt, persona n°2, diagramme de cas d'utilisation, d'activité, Trello, maquette affichage profil• Scraping d'un document au format PDF• Récupération et affichage des informations des candidats sur angular• Rédaction du rapport et du diaporama
PLANCHENAU Simon	<ul style="list-style-type: none">• Génie Logiciel : backlog, maquettes du projet• Implémentation d'un fichier JSON dans le BDD• Création et développement entier de l'application web avec Angular• Gestion des filtres (requêtes avec le serveur et la BDD)
PROUDY Julien	<ul style="list-style-type: none">• Génie Logiciel : Gantt, sprint, backlog, persona n°1, diagramme de séquence (version 1), Trello• Mise en place de la BDD et du serveur• Rédaction document "Traitement des données issues du scraping Pôle Emploi"• Codage de l'indicateur (radar chart)

TABLE 1 – Répartition des tâches

6.2 Outils utilisés

De nombreux outils web de collaboration, communication et de génie logiciel ont été utilisés au cours de la réalisation du projet :

- **Google Drive** pour la rédaction des comptes-rendu de chaque séance de travail, le Gantt, document de reflexion sur le projet, etc.
- **Trello** pour la création d'un tableau Kanban afin d'effectuer le suivi du projet et sa progression.
- Un **Gantt** afin de faire un macro-planning du projet.
- **Overleaf** pour la rédaction du rapport en $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$.
- **Github** afin de stocker le projet et de le partager entre les membres de l'équipe.
- **Uizard** pour la réalisation des maquettes de l'application web.
- **Terrastruct** pour la réalisation de l'architecture logicielle.
- **PowerPoint** pour la création du diaporama.
- **VisualStudio Code & Suite JetBrains** environnements de développement.

7 Réalisation

Le résultat de l'application HRhelper est en accord avec ce qui été souhaité. Les diagrammes de cas d'utilisation et d'activité correspondent à l'état actuel du projet. Le scraping est fonctionnel, les données sont récupérées et stockées de façon sécurisée et le système de filtres est opérationnel avec la BDD qui s'actualise en fonction de ce dernier. De plus, l'affichage des profils se fait dynamiquement et la navigation sur l'application web est intuitive.

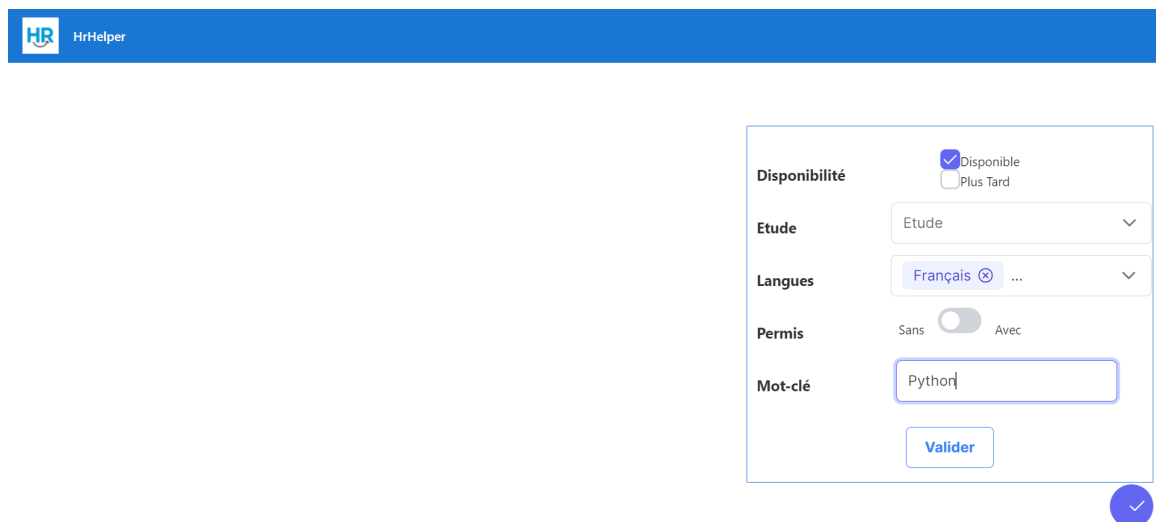


FIGURE 11 – Affichage des filtres dans HRhelper

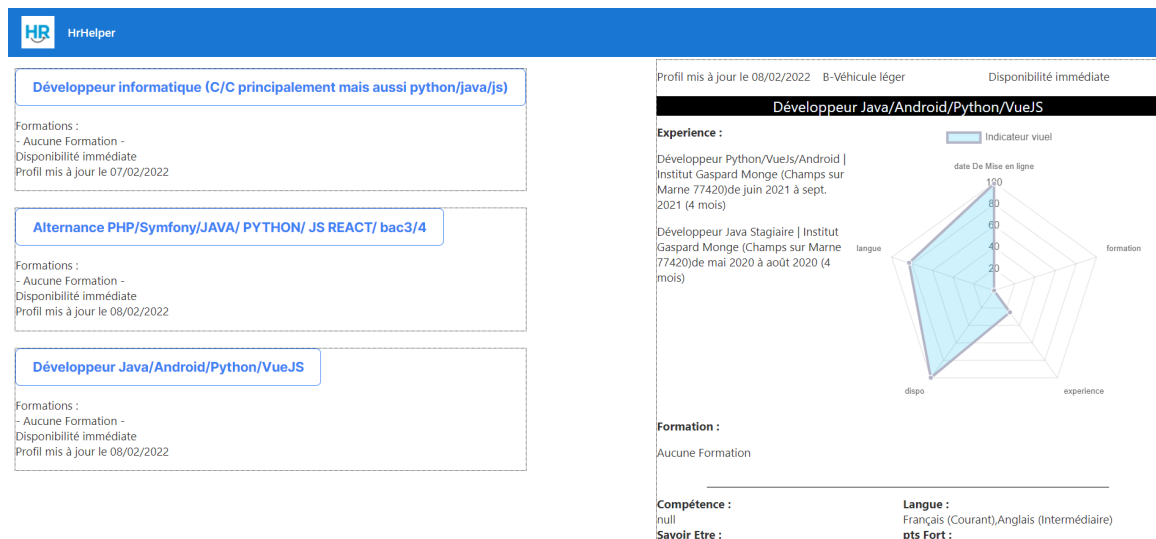


FIGURE 12 – Affichage d'un profil sur l'application web

8 Conclusion

Ce projet a permis à l'équipe d'apprendre à maîtriser différentes technologies telles que le scraping, Angular, MongoDB, NodeJS mais aussi la gestion de projet de manière Agile avec la méthode SCRUM. Nous avons également assimilé l'importance du client et de ses attentes dans ce type de projet, de même pour l'utilisation et le traitement des données dites "sensibles" du fait que ça soit des informations personnelles.

Grâce à HRhelper les recruteurs et responsables RH ont un outil opérationnel, efficace et innovant. La conception de l'application est généraliste dans son code et permet donc d'être réutiliser dans d'autres domaines comme par exemple pour la comparaison des prix pour un même produit.

Des améliorations auraient pu être apportées si nous avions eu du temps supplémentaire. En effet nous aurions pu faire en sorte que l'utilisateur puisse contacter le ou les candidats repérés, ou encore faire une page "A propos", rendre le site responsive avec l'optimisation pour téléphone, permettre à un utilisateur d'avoir un compte mais aussi de rechercher des profils dans une base de données plus complète (lancement d'un nouveau scraping).

9 Annexes

9.1 Traitement des données issus du scraping du site Pôle Emploi

Le but de ce traitement des données est d'obtenir des indicateurs uniques à chaque profil dans le but d'avoir un graphique associé à ce profil qui permettra à l'utilisateur d'HRHelper d'avoir un visuel en fonction de sa recherche

Cet élément va avoir un rôle déterminant dans le choix de la sélection de profil de l'utilisateur, il faut donc que les indicateurs soient faits de manière à ce que cela ne soit pas discriminatoire et surtout qu'ils soient personnalisés --en fonction de la recherche.

Nous allons avoir 3 indicateurs dont voici les domaines :

- Le niveau en langue vivante
- Les expériences professionnelles
- Les formations

1. L'indicateur du niveau en langue vivante

Le calcul de cet indicateur se fera en fonction de la nature de la recherche de l'utilisateur, en effet certains corps de métier nécessitent une connaissance plus ou moins accrue d'une langue vivante supplémentaire.

Par exemple, si un RH recherche un développeur Python, l'indicateur de langue doit être élevé si un profil a un très bon niveau en anglais. Au contraire, si la recherche porte sur un préparateur de commande, le niveau en anglais est anecdotique donc si le profil maîtrise moyennement l'anglais, l'indicateur sera élevé également.

Nous pouvons également inclure le niveau en langue française pour les travailleurs d'origine étrangère mais cela pourrait potentiellement poser des problèmes étant donné que certains profils ne renseignent pas leur niveau dans leur langue maternelle.

2. L'indicateur des expériences professionnels

Cet indicateur est l'un des plus important, il sera calculé en fonction de la recherche de l'utilisateur et des expériences du profil (la nature de l'expérience et la durée de cette dernière).

L'indicateur sera défini avec plusieurs critères évolutifs, en fonction de la recherche. En effet, une recherche avec un mot-clé, une année d'expérience voulut va prendre plus de critère qu'une recherche basique avec seulement un mot-clé.

A partir d'une recherche, les critères seront alors fixés et nous pourrons calculer notre indicateur :

$$\text{- Critère 1} = \frac{\text{Année d'expérience totale}}{\text{Année d'expérience souhaitée}}$$

Le nombre d'années d'expérience du profil sera calculé en fonction de la recherche pour ne prendre que les expériences en rapport avec le métier recherché.

$$\text{- Critère 2} = \frac{\text{Année d'expérience totale}}{\text{Année d'expérience souhaitée}} \times \text{multiplicateur d'entreprise}$$

Le multiplicateur d'entreprise sera appliqué si une entreprise reconnue est indiquée dans les expériences du profil. Ce multiplicateur sera plus ou moins grand en fonction de l'entreprise.

$$\text{- Critère 3} = \frac{\text{nombre de mots clé trouvé}}{\text{nombre de mots clés total}}$$

Ici, nous prenons en compte les mots clés renseignés par l'utilisateur. Plus le nombre de mots clés contenu dans le profil correspond à la recherche, plus l'indicateur sera élevé.

Après avoir calculé les critères en fonction de la recherche de l'utilisateur, il suffit de les multiplier entre eux afin d'obtenir un indice final qui sera affiché sur le graphique.

Le nombre de critères, ainsi que les critères actuels seront susceptibles d'évoluer.

3. L'indicateur des formations

L'indicateur des formations sera en fonction de la recherche de l'utilisateur. On peut alors déduire 3 critères :

$$\text{- Critère 1} = \frac{\text{Niveau après le bac du profil}}{\text{Niveau après le bac voulu}}$$

Si le profil recherché doit être de bac +5, alors plus le profil se rapproche de ce niveau, plus l'indicateur sera élevé.

$$\text{- Critère 2} = \frac{\text{Niveau après le bac du profil dans le domaine}}{\text{Niveau après le bac voulu dans le domaine}}$$

Si le profil recherché doit être de bac +5 dans le domaine de l'informatique, alors plus le profil se rapproche de ce niveau, plus l'indicateur sera élevé.

$$\text{- Critère 3} = \frac{\text{Niveau après le bac du profil dans le domaine}}{\text{Niveau après le bac voulu dans le domaine}} \times \text{indice formation}$$

Indice de la formation sera en fonction de la formation du profil, si c'est un master, un DUT, une école d'ingénieur.

9.2 Fiche d'activité DUBIN Baptiste

Semaine	du	au	Tâches	duree	Organisation	
44	01/11/21	07/11/21	Réunion n°1	01:30:00	Toute l'équipe	Gestion Projet
45	08/11/21	14/11/21	Rédaction et recherche des idées projet	02:00:00	Toute l'équipe	Gestion Projet
46	15/11/21	21/11/21	Recherche des idées / découverte du scraping	03:00:00	Toute l'équipe	Gestion Projet
47	22/11/21	28/11/21	Diagramme de Classe/ Début scraping data	03:00:00	Tout seul/ Tout seul	Gestion Projet
48	29/11/21	05/12/21	Se logger sur LinkedIn et naviguer de profil en profil	03:00:00	seul	Récupération Données
49	06/12/21	12/12/21				
50	13/12/21	19/12/21				
51	20/12/21	26/12/21				
52	27/12/21	02/01/22				
1	03/01/22	09/01/22	Scraping page linkedin DONE/ réunion avec l'équipe	03:00:00	seul/ Toute l'équipe	Récupération Données
2	10/01/22	16/01/22	Début scraping pôle emploi/ Prise en main selenium	03:00:00	Seul	Récupération Données
3	17/01/22	23/01/22	Scraping Pole Emploi DONE/Fix Problème JSON/ Test linkedin recruter account pour scraper plusieurs profils	03:00:00 03:00:00	Seul	Récupération Données/Maintenance
4	24/01/22	30/01/22	Début filtrage pour réduire taille JSON/ Début création matrice pour radar charts et mise en place des formules	03:00:00	Julien/seul	Optimisation
5	31/01/22	06/02/22	Correction de bug/ regroupement des fichiers des fichiers python	03:00:00	Seul	Optimisation
6	07/02/22	13/02/22	Finalisation des connections entre fichiers python/Correction des erreurs relative connection a pôle emploi	03:00:00	Seul	Optimisation
7	14/02/22	20/02/22	Correction de bugs/ préparation pre-soutenance	03:00:00	Groupe	Préparation soutenance
8	21/02/22	27/02/22				
9	28/02/22	06/03/22				
10	07/03/22	13/03/22				
11	14/03/22	20/03/22				
12	21/03/22	27/03/22				
13	28/03/22	03/04/22				

FIGURE 13 – Fiche de suivi d'activité de Baptiste DUBIN

9.3 Fiche d'activité MOTTIER Anaïs

Semaine	du	au	Tâches	Durée	Organisation	
44	01/11/21	07/11/21	Réunion n°1 : rencontre avec le client	01:30:00	Toute l'équipe	Gestion Projet
45	08/11/21	14/11/21	Rédaction et recherche des idées pour le projet	02:00:00	Toute l'équipe	Analyse et Conception
46	15/11/21	21/11/21	Reprendre les points abordés lors de la 1ère réunion	03:30:00	Toute l'équipe	Gestion Projet
			Création des backlog, des sprints et mise à jour du planning			
			Renseignement sur la méthode Agile et Scrum			
			Rendez-vous avec le client			
47	22/11/21	28/11/21	Trouver plateforme alternative à LinkedIn, CVthèque	02:00:00	Seul	Analyse et Conception
			Diagramme de cas d'utilisation (version 1)	01:00:00		
48	29/12/21	05/12/21	Prise en main des technologies et librairies de scraping.	03:00:00	Seule	Traitement de données
49	06/12/21	12/12/21	PARTIELS & VACANCES			
50	13/12/21	19/12/21				
51	20/12/21	26/12/21				
52	27/12/21	02/01/22				
1	03/01/22	09/01/22	Lecture d'un fichier pdf	03:00:00	Seule	Traitement de données
2	10/01/22	16/01/22	Lecture fichier pdf + transformation en format Json	06:00:00	Seule	Traitement de données
3	17/01/22	23/01/22	Transformation en format Json -> Fini	05:00:00	Seule	Traitement de données
			Récupération des infos d'un profil pour la liste des profils (display.components)	03:00:00	Seule	Web
4	24/01/22	30/01/22	Affichage dynamiques des profils (itérations & liaison avec le SCSS)	11:30:00	Seule et avec Simon	Web
			Masquer liste profil quand les filtres n'ont pas été renseignés	01:30:00	Seule	Web
5	31/01/22	06/02/22	Rédaction du rapport + réunion avec le client	03:00:00	Seule	Gestion Projet
			Rédaction rapport + UML + gestion de projet	17:00:00	Seule	Gestion Projet
6	07/02/22	13/02/22	Architecture logicielle	00:30:00	Avec Simon	Analyse et Conception
			Correction orthographe site web	00:05:00	Seule	Web
7	14/02/22	20/02/22	Lecture collective rapport et dernières modifications	02:00:00	Toute l'équipe	Gestion Projet
			Création du diaporama	05:00:00	Seule	Gestion Projet
			Correction rapport suite au retour du tuteur	02:00:00	Seule	Gestion Projet
			Préparation oral soutenance	00:15:00	Toute l'équipe	Gestion Projet
			Soutenance blanche avec Madeth MAY (18/02 à 14h)	01:00:00	Toute l'équipe	Gestion Projet
8	21/02/22	27/02/22	Pré-soutenance 22/02 à 15h45	00:30:00	Toute l'équipe	Gestion Projet
9	28/02/22	06/03/22				
10	07/03/22	13/03/22				
11	14/03/22	20/03/22				
12	21/03/22	27/03/22				
13	28/03/22	03/04/22				

FIGURE 14 – Fiche de suivie d'activité de Anaïs MOTTIER

9.4 Fiche d'activité PLANCHENAULT Simon

Semaine	du	au	Tâches	duree	Organisation	
44	01/11/21	07/11/21	Réunion n°1	01:30:00	Toute l'équipe	Gestion Projet
45	08/11/21	14/11/21	Rédaction et recherche des idées projet	02:00:00	Toute l'équipe	Gestion Projet
46	15/11/21	21/11/21	Recherche des idées Découverte du scraping	03:00:00	Toute l'équipe	Gestion Projet
47	22/11/21	28/11/21	Création maquette du site Découverte scraping et gspread	03:00:00	seul	Gestion de Projet Traitement de données
48	29/12/21	05/12/21	gspread / panda	03:00:00	seul	Traitement de données
49	06/12/21	12/12/21				
50	13/12/21	19/12/21				
51	20/12/21	26/12/21				
52	27/12/21	02/01/22				
1	03/01/22	09/01/22	Implémentation du fichier json dans le Google sheet Réunion avec l'équipe Suite de la creation de la page web	06:00:00	seul / Baptiste	Traitement de données
2	10/01/22	16/01/22	Implémentation du fichier json dans MangoDB Ajout d'un template bootstrap (abandon) Continuer sur l'angular en ajoutant les modules ainsi que les interactions dans les modules et ajout de librairie d'interface Maquettes	09:00:00	seul	Interface
3	19/01/22	23/01/22	Création et interactions des boutons du filtre	10:00:00	seul	Interface
4	24/01/22	30/01/22	Création du component information- personne Début de la création de la méthode post sur la base de données	09:00:00	seul	Interface Techno
5	31/01/22	06/02/22	Création de l'update Ajout du component d'affichage Réorganisation de l'angular Css Ajout des graphiques	11:00:00	seul	Interface Techno
6	07/02/22	13/02/22	Ajustage css Début du rapport Exécution d'un programme python dans angular Documentation du code	08:00:00	seul / Anaïs	Interface Redaction
7	14/02/22	20/02/22	Redaction rapport Correction de bugs sur Angular / préparation pre-soutenance	05:00:00	seul/Anaïs	redaction / interface
8	21/02/22	27/02/22				
9	28/02/22	06/03/22				
10	07/03/22	13/03/22				
11	14/03/22	20/03/22				
12	21/03/22	27/03/22				
13	28/03/22	03/04/22				

FIGURE 15 – Fiche de suivi d'activité de Simon PLANCHENAULT

9.5 Fiche d'activité PROUDY Julien

Semaine	du	au	Tâches	duree	Organisation	
44	01/11/21	07/11/21	Réunion n°1	01:30:00	Toute l'équipe	Gestion Projet
45	08/11/21	14/11/21	Rédaction et recherche des idées projet	02:00:00	Toute l'équipe	Gestion Projet
46	15/11/21	21/11/21	Réunion avec le chef de projet (tuteur) pour validé notre vision du projet et poser certaines questions Validation du planning + Sprint / backlog (possibilité d'évolution)	03:00:00	Toute l'équipe Toute l'équipe	Gestion Projet
47	22/11/21	28/11/21	Projection dans le planning Réalisation diagramme de séquence Création d'un persona RH	03:00:00	Toute l'équipe Seul Seul	Gestion Projet
48	29/11/21	05/12/21	Prise en main de filtre sur un dataset en Python Mise a jour du Trello en format Sprint	03:00:00	Seul Seul	Gestion Projet Techno
49	06/12/21	12/12/21				
50	13/12/21	19/12/21				
51	20/12/21	26/12/21				
52	27/12/21	02/01/22				
1	03/01/2022	09/01/22	Réunion de groupe pour repartition du travail Affichage du json obtenu par scrapping sur un page web	03:00:00	Toute l'équipe Seul	Gestion Projet Techno
2	10/01/22	16/01/22	Utilisation de l'API Google Sheets pour relier la BDD à Angular -> abandon BDD désormais sur MongoDB mise en place d'un Back-End Node JS pour transmettre les données sur Angular (Mais problème de connexion avec MongoDB)	04:00:00	Seul	Techno
3	17/01/22	23/01/22	[18/01] Connexion entre Back-End et Angular, le transfert de donnée fonctionne mais problème avec CORS [20/01] Problème avec CORS résolu avec la mise en place d'un proxy, la connexion avec la BDD est opérationnelle	03:00:00 03:00:00	Seul	Techno
4	24/01/22	30/01/22	[25/01] Début filtrage pour réduire taille JSON/ Début création matrice pour radar charts et mise en place des formules [27/01] Ajout des indices uniques a chaque profils en fonction du CV (pour faire les charts)	03:00:00 03:00:00	Julien / Baptiste Seul	Techno
5	31/01/22	06/02/22	[03/02] Finalisation des indicateurs et reflexion sur une version des indicateurs moins discriminatoire avec le tuteur	03:00:00	Seul	Techno
6	07/02/22	13/02/22	[10/02] Rédaction d'un document explicatif en fonction des recommandations du tuteur (explication des calculs ainsi que le raisonnement derrière chaque indicateurs)	03:00:00	Seul	Techno
7	14/02/22	20/02/22	Preparation pré-soutenance (organisation prise de parole et suggestion diapo / rapport) [19/02] Soutenance blanche a distance avec le tuteur	03:00:00	Toute l'équipe	Gestion Projet
8	21/02/22	27/02/22	Pré-soutenance 22/02 à 15h45			
9	28/02/22	06/03/22				
10	07/03/22	13/03/22				
11	14/03/22	20/03/22				
12	21/03/22	27/03/22				
13	28/03/22	03/04/22				

FIGURE 16 – Fiche de suivie d'activité de Julien PROUDY

Lexique

Backlog liste de fonctionnalités d'un projet sur lesquelles du travail reste à faire. 13

CAPTCHAs mesure de sécurité de type "authentification par question-réponse" afin de s'assurer qu'une personne réelle est bien derrière l'écran et non un robot. 7

cas d'utilisation différentes manière d'utiliser un produit, ensemble d'actions possibles.. 3, 4

component classe, élément pouvant être réutiliser dans une application. Le component expose une vue et définit la manière dont l'utilisateur peut interagir avec. 12

CSV (Comma-Separated Values) fichier informatique de type tableur. 7

CVthèque plateforme en ligne sur laquelle on trouve un ensemble de CV que les candidats à un emploi ont eux-mêmes déposer. 8, 11

daily metting réunion courte et quotidienne qui rassemble l'ensemble des membres d'une équipe travaillant sur un projet. Elle a pour but de faire le point rapidement et efficacement sur l'avancé du projet. 13

framework infrastructure logicielle désignant un ensemble d'outils et de composants logiciels à la base d'un logiciel ou d'une application. 11

JSON (JavaScript Objet Notation) type de fichier permettant de représenter des données structurées dans un langage léger. 8, 11

macro-planning planning permettant de donner une visibilité globale sur l'organisation d'un projet d'un point de vue temporel.. 13

page Honeyspot méthode de défense active consistant à attirer des robots et programmes informatiques sur des ressources de type "leurre" afin de les identifier et de les neutraliser. 7

persona personne imaginaire représentant un groupe cible dans le cadre du développement d'un nouveau projet. 3

protocole HTTP (Hypertext Transfer Protocol) protocole de communication entre un client et un server informatique. Il permet à un utilisateur d'accéder au contenu d'une page Web via un navigateur. 7

radar chart aussi appelé diagramme de Kiviat, diagramme en radar, en radar, en étoile ou en toile d'araignée, ce graphique sert à représenter en deux dimensions au moins 3 données. 8, 9, 12

SCRUM méthode agile dans le domaine de la gestion de projet ayant pour objectif d'améliorer la productivité des équipes travaillant sur le projet concerné. 13

sprint phase définit temporellement permettant de décomposer un processus de développement d'un projet. 13

Users Stories phrases simples rédigés dans un langage courant permettant de décrire les fonctionnalités d'un projet à développer. 3, 6

XML (Extensible Markup Language) est un langage informatique utilisé pour la conception de sites web. 7

éthique ensemble de principes moraux et valeurs à respecter. 3, 6, 7