

UFR Sciences et Techniques – Département Informatique

Compte-rendu de la réunion 6

"Projet M1 - HR-Helper"

Mardi 11 janvier 2022

Présents : Julien PROUDY, Simon PLANCHENAU, Anaïs MOTTIER, Baptiste DUBIN

Points abordés :

1. Scraping de Pôle emploi
2. Interface Angular
3. API Google Sheet
4. Scraping d'un CV en pdf + récupération sous format Json

Rédacteur / Relecteur : Anaïs MOTTIER - Simon PLANCHENAU - Baptiste DUBIN - Julien PROUDY

Début de séance : 13h30 **Fin de séance :** 16h30

Problème avec linkedin même en utilisant un compte recruteur (payant), on ne peut avoir uniquement accès au profil des personnes de notre réseau.

1 - Scraping de Pôle emploi

Commencement scrapping pôle emploi, utilisation de scrapy impossible sur ce site utilisation de selenium malgré que cela soit plus lent.

2 - Interface Angular

Créations de composants avec l'ajout de fonctionnalités (bouton, zone de texte). Interaction avec les boutons dans le composants filtre.

3 - API Google Sheet

[En cours] Intégration de l'API Google Sheet dans un composant Angular.

4 - Scraping d'un CV en pdf + récupération sous format Json.

Possibilité de lire un fichier pdf terminé. La récupération des données sous format Json est en cours de développement.

Prochaine réunion : Mardi 18 janvier (9h00 - 12h00)

Compte-rendu de la réunion 5

"Projet M1 - HR-Helper"

Mardi 4 janvier 2022

Présents : Julien PROUDY, Simon PLANCHENAULT, Anaïs MOTTIER, Baptiste DUBIN

Points abordés :

5. Scraping de LinkedIn.
6. Récupération du scraping et incrémentation de la base de données (google sheets).
7. Récupération des données de la BDD et affichage sur une page web.
8. Scraping d'un CV en pdf.

Rédacteur / Relecteur : Anaïs MOTTIER - Simon PLANCHENAULT - Baptiste DUBIN - Julien PROUDY

Début de séance : 9h00 **Fin de séance** : 12h00

1 - Scraping de LinkedIn.

Scraping terminé pour une personne donnée, code déposé et documenté à 50% (pas avec pydoc).

2 - Récupération du scraping et incrémentation de la base de données (google sheets).

Récupération du fichier json du scraping pour le filtrer et l'ajouter sur le google sheet afin avoir une base de données propre et exploitable.

3 - Récupération des données de la BDD et affichage sur une page web.

Récupération des données depuis le .json obtenu suite au scraping d'un compte LinkedIn. On affiche les données de ce fichier sur la console d'une page web (depuis un serveur python). Mais, par la suite nous allons utiliser une BDD sous sheet pour le traitement des données.

4 - Scraping d'un CV en pdf.

Travail sur la lecture d'un CV en pdf afin de pouvoir récupérer les informations importantes pour par la suite les utiliser dans des filtres. Technique pas encore opérationnelle en fin de séance.

Prochaine réunion : Mardi 11 janvier (13h30 - 16h30)

Compte-rendu de la réunion 4

"Projet M1 - HR-Helper"

Jeudi 2 décembre 2021

Présents : Julien PROUDY, Simon PLANCHENAULT, Anaïs MOTTIER, Baptiste DUBIN

Points abordés :

1. Réorganisation du Trello.
2. Prise en main des technologies et librairies de scraping.
3. Suite de la prise en main de gspread et panda.
4. Prise en main des filtres.

Rédacteur : Anaïs MOTTIER **Relecteur** : Simon PLANCHENAULT - Baptiste DUBIN - Julien PROUDY

Début de séance : 9h00 **Fin de séance** : 12h00

1 - Réorganisation du [trello](#).

Après avoir envoyé notre Trello à Claudine Piau-Toffolon, elle nous a fait un retour sur celui-ci ce qui entraîne des changements.

2 - Prise en main des technologies et librairies de scraping.

Pour récupérer les données nous allons utiliser les librairies Beautiful Soup et Scrapy présentes en Python.

3 - Suite de la prise en main de gspread et panda.

Exportation et importation de données de google sheets depuis python. Ainsi que la prise en main de panda pour trier nos données.

4 - Prise en main des filtres.

Réalisation de différents filtres à partir d'un dataset en Python.

Prochaine réunion : Mardi 4 janvier (9h00 - 12h00)

Compte-rendu de la réunion 3

"Projet M1 - HR-Helper"

Mardi 23 novembre 2021

Présents : Julien PROUDY, Simon PLANCHENAULT, Anaïs MOTTIER, Baptiste DUBIN

Points abordés :

1. Compléter le planning.
2. Trouver une alternative pour LinkedIn.
3. Réalisation d'un début de maquette.
4. Réalisation des diagrammes de classe, séquence et cas d'utilisation.
5. Prise en main de Gspread.
6. Création du persona RH

Rédacteur : Anaïs MOTTIER **Relecteur** : Simon PLANCHENAULT - Baptiste DUBIN - Julien PROUDY

Début de séance : 13h30

Fin de séance : 16h30

1 - Compléter le [planning](#).

Compléter le planning en prévoyant les tâches à des dates données.

2 - Trouver une alternative pour LinkedIn.

Durant le week-end, nous avons appris que le site de LinkedIn interdisait le scraping sur son site. Nous essayons donc de trouver une autre plateforme du même type sur laquelle on puisse récupérer des informations sur les candidats à l'emploi.

-> Possibilité de trouver des cv et profils sur : [Pôle emploi](#), google, [emploiLibre](#) (besoin d'un compte)

=> Finalement, on peut scaper LinkedIn avec l'utilisation d'un VPN.

3 - Réalisation d'un début de maquette.

The wireframe shows a web browser window with the address 'https://hrHelper'. The page title is 'HrHelper'. On the left, there is a sidebar with a 'filtre' dropdown menu (options 1, 2, 3, 4), a 'Checkbox', and a 'Liste des filtres :'. The main content area has a search bar with 'Intitulé du metier :', 'Lieux :', and a 'Rechercher' button. Below the search bar, there are two identical job listing cards. Each card contains a text input for 'Intitulé du metier', a text area for 'Texte descriptif de la personne', and a text area for 'Les points fort :'. The browser window also shows standard navigation icons and a search bar.

4 - Réalisation des [diagrammes de classe, séquence et cas d'utilisation](#).

5 - Prise en main de Gspread.

Gspread permet de placer les données récupérées dans un google sheet.

6 - Création du persona RH.

 Nom Pauline Intitulé de poste Recruteur (RH) Âge 29 ans Niveau d'études Bac +3 Réseaux sociaux      Secteur d'activité Assurance	Objectifs Trouver les meilleurs profils possible pour un poste par rapport à un ensemble de candidatures	Outils nécessaires au quotidien <ul style="list-style-type: none"> • Outil d'e-mailing • Logiciel de planning des salariés • Logiciel de traitement de texte • Système de gestion de contenu
	Principaux défis <ul style="list-style-type: none"> • Motivation des salariés • Communication interne • Gestion des changements • Ressources 	Indicateurs de performance En charge de la gestion des employés de l'entreprise
	Moyen de communication préféré <ul style="list-style-type: none"> • E-mail • Téléphone • Face à face • Réseaux sociaux • • 	Sources d'information Formations Responsabilités Recruter des bons éléments pour l'entreprise Supérieur hiérarchique DRH

Prochaine réunion : Lundi 15 novembre 2021 (13h00-16h00)

Compte-rendu de la réunion 2

"Projet M1 - HR-Helper"

Lundi 15 novembre 2021

Présents : Julien PROUDY, Simon PLANCHENAULT, Anaïs MOTTIER, Baptiste DUBIN

Points abordés :

1. Reprendre les points abordés lors de la 1ère réunion
2. Création des backlog, des sprints et mise à jour du planning
3. Renseignement sur la méthode Agile et Scrum
4. Documentation et prise en main du scraping

Rédacteur : Anaïs MOTTIER - Julien PROUDY

Relecteur : Simon PLANCHENAULT - Baptiste DUBIN

Début de séance : 13h00

Fin de séance : 16h00

1 - Reprendre les points abordés lors de la première réunion.

https://docs.google.com/document/d/1p0GZRNlvc9Kd6WZzv0YpYm3dd_lvrB7W6dPwDfL-Fyl/edit?usp=sharing

2 - Création des backlog, des sprints et mise à jour du [planning](#).

Commencer à lister les spécificités et les étapes du projet.

3 - Renseignement sur la méthode Agile et Scrum.

Lecture de la documentation trouvée sur la page Umtice du module et sur internet.

4 - Documentation et prise en main du scraping.

Test de scraping sur le site de LinkedIn avec scrapy.

Prochaine réunion : Mardi 23 novembre 2021 (13h30 - 16h30)

UFR Sciences et Techniques – Département Informatique

Compte-rendu de la réunion 1

"Projet M1 - HR-Helper"

Jeudi 4 novembre 2021

Présents (4 personnes) : Julien PROUDY, Simon PLANCHENAULT, Anaïs MOTTIER, Baptiste DUBIN

Points abordés : Rencontre avec le client, définitions de ses attentes et de l'organisation.

Rédacteur : Anaïs MOTTIER - Julien PROUDY

Relecteur : Simon PLANCHENAULT - Baptiste DUBIN

Début de séance : 13h30

Fin de séance : 16h30

Prochaine réunion : Lundi 15 novembre 2021 (13h00-16h00)