

RAPPORT DE STAGE

Rapport de stage effectué 24/06/2024 au 19/07/2024

Tuteur de stage : LEMEE Christophe

Professeur référent : MORVAN Isabelle

Je soussigné, Marin Vandelet, étudiant en 1ère année de préparation au Bachelor MMI, certifie sur l'honneur que les travaux soumis en mon nom sont le fruit de mes efforts et réflexions personnelles.

Toute idée ou tout document utilisé pour étayer ce travail et ne constituant pas une réflexion personnelle est noté entre guillemets et il est précisé la source dont il est extrait, le nom de l'auteur et la référence du document. Signature:



1. Présentation Générale de l'Entreprise Deepblock

1.1 Le Secteur d'Activité et place de l'entreprise dans le secteur

Crée en 2017 par Christophe LEMÉE (Ingénieur) et Maître Jean-Philippe TOUATI (Avocat au Barreau de Paris), Deep block est la fusion du Deep Learning et de Blockchain. La société est une Legaltech spécialisée dans la transformation numérique et la réalisation de système d'information sur-mesure pour les entreprises du monde juridique, mais pas exclusivement.

Elle a aussi depuis 7 ans maintenant, conçu de nombreux modules fonctionnels, en lien avec son ADN de transparence et de souveraineté numérique tel que le dépôt de capital, de multiples certificats à usages différents et même de l'IA. Plusieurs de ses modules sont accessibles par API et sont consommés par des éditeurs de plateformes, mais aussi par des entreprises des secteurs Privées comme Publiques. Concrètement, Deep Block évolue dans le secteur de l'intelligence artificielle (IA) et des technologies numériques. Ce secteur regroupe un ensemble d'activités liées à la conception et à l'implémentation de systèmes intelligents, capables de traiter des volumes massifs de données, de prendre des décisions automatisées et d'exécuter des tâches complexes. Ces technologies sont particulièrement prisées dans divers secteurs, comme la finance, la logistique, l'industrie et la santé, où elles apportent des innovations majeures. L'entreprise possède plusieurs sites en île de France ou les employés travaillent en collaboration que ce soit physiquement ou à distance. Il y a notamment une équipe de développeurs qui travaille pour l'entreprise.

Bien que jeune, l'entreprise s'est rapidement imposée comme un acteur clé dans le secteur des technologies de l'intelligence artificielle et de la blockchain. Grâce à son expertise pointue, Deepblock a su développer des solutions innovantes qui répondent aux besoins des entreprises en matière de gestion de données et de sécurité des transactions. L'un des atouts majeurs de Deepblock est son service de signature électronique, qui se distingue par plusieurs valeurs ajoutées. Deep Block a été une des premières entreprises à proposer ce genre de service. Il est à savoir que tous les services que propose deep block sont disponibles sur leur site en ligne.

La signature électronique de Deepblock utilise la technologie blockchain pour garantir une sécurité maximale, assurant que chaque signature est infalsifiable et que l'intégrité des documents est maintenue. Cette approche permet non seulement de sécuriser les transactions, mais aussi de tracer et d'authentifier chaque étape du processus, offrant une transparence totale. De plus, ce service est conçu pour être facilement intégrable aux systèmes existants des entreprises, simplifiant ainsi la transition vers des processus numériques. La signature électronique de Deepblock est conforme aux réglementations les plus strictes, telles que le règlement eIDAS en Europe, ce qui assure sa validité légale à l'échelle internationale. Ce service permet aussi de réduire les coûts liés à la gestion des documents physiques, d'accélérer les processus de validation et de faciliter les collaborations à distance, tout en améliorant l'expérience utilisateur grâce à une interface intuitive et accessible.

1.2 Sa Particularité

Ce qui fait la particularité de Deepblock, c'est son approche multidimensionnelle combinant intelligence artificielle, big data et blockchain. Contrairement à beaucoup d'autres acteurs du marché qui se concentrent sur une seule technologie, Deepblock propose des solutions intégrées permettant de couvrir un spectre plus large de besoins métiers. L'entreprise se distingue également par son fort engagement en matière de recherche et développement. En investissant constamment dans l'innovation, Deepblock reste à la pointe des dernières avancées technologiques, garantissant ainsi à ses clients des solutions toujours à jour et optimisées pour leurs besoins spécifiques. Deepblock a reçu de nombreuses récompenses prestigieuses, notamment un Lauréat Palmarès du droit « Signature Électronique » 2022 , le label "Entreprise Innovante" décerné par Bpifrance en 2022, et a été classé parmi les "Top 100 Startups" par French Tech en 2023. Ces récompenses témoignent de l'innovation dont fait preuve l'entreprise la plupart du temps.



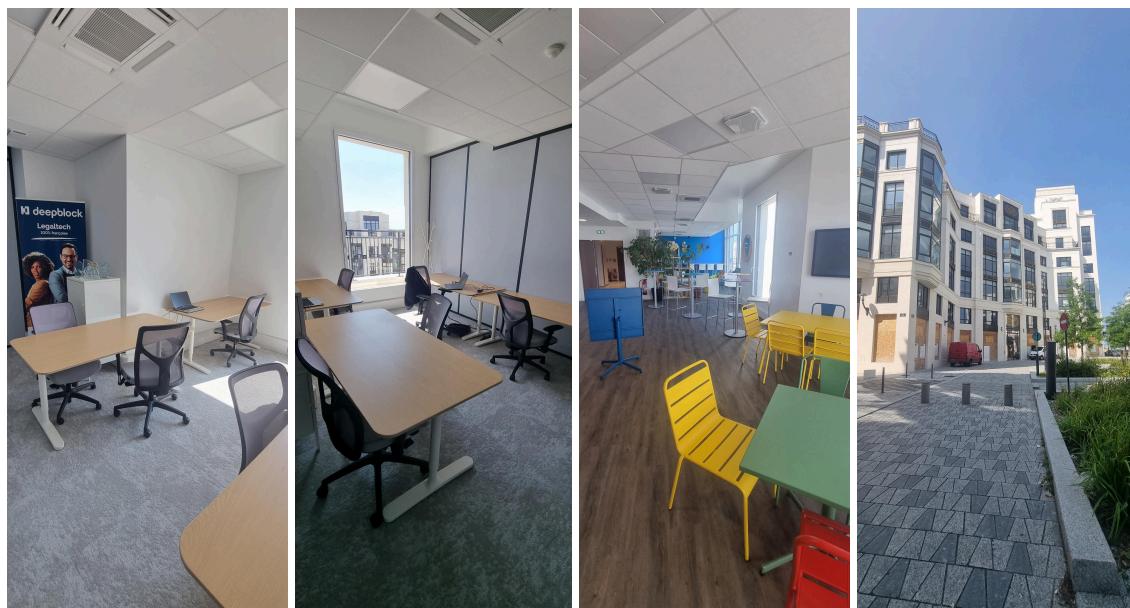
1.3 Chiffre d'Affaires

Le chiffre d'affaires de Deepblock est une bonne indication et reflète sa croissance et son succès, et ce dans un marché hautement prisé. En 2022, l'entreprise a enregistré un chiffre d'affaires d'environ 2 millions d'euros, ce qui montre une croissance depuis sa création. Ce chiffre reflète aussi la confiance des clients dans les solutions proposées par Deepblock.

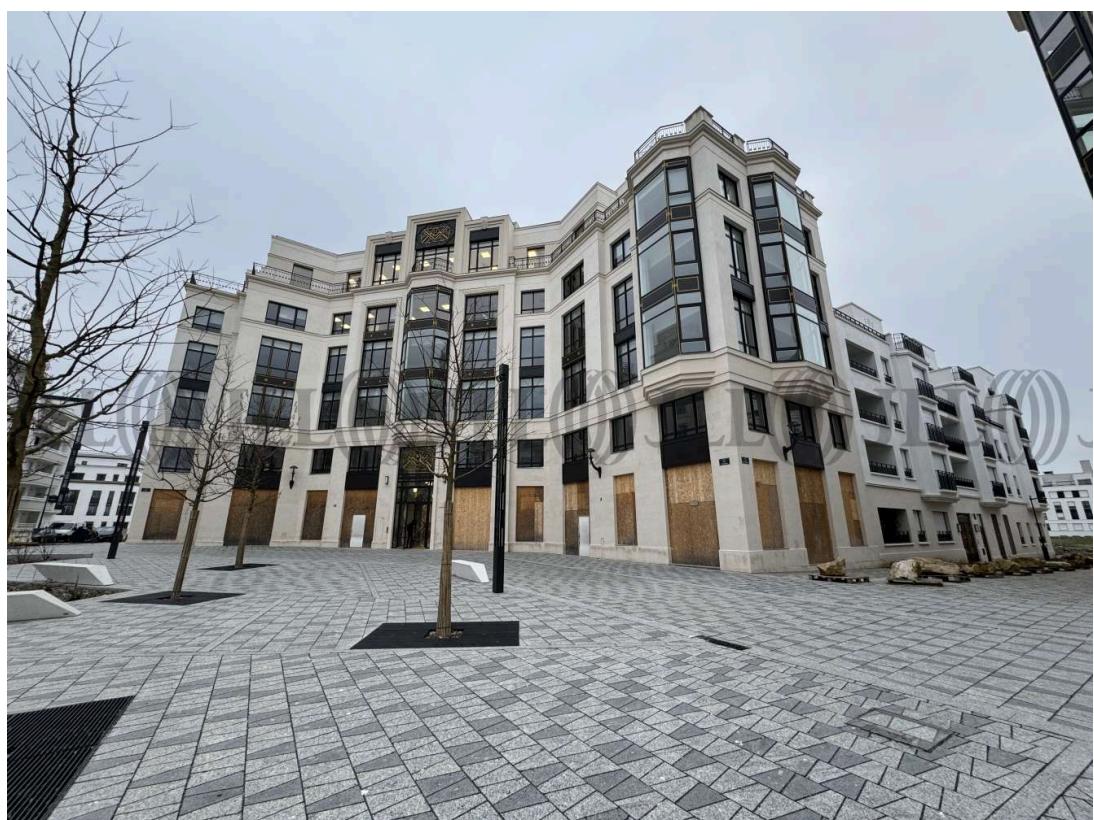
1.4 Contexte du stage dans l'entreprise

J'étais sur le site de Chessy, place octogonale dans un bureau géré par Village by CA, un réseau d'accélérateurs de start-ups créé par le Crédit Agricole pour soutenir l'innovation et l'entrepreneuriat. Le site de Chessy, situé sur une place octogonale, est un espace moderne où les start-ups comme Deep Block peuvent bénéficier d'un environnement dynamique, de ressources partagées, et d'un réseau d'entreprises et de partenaires pour favoriser la

croissance et le développement. L'étage où nous travaillons était agréable, spacieux avec un coin cuisine ou encore une grande salle de repos. Nous avions des codes d'accès wifi temporaire et nous travaillons avec nos pc portables pour permettre de travailler à distance aussi.



Il faut savoir que je travaillais la moitié du temps en présentiel et l'autre en télétravail car toutes les tâches dont je devais m'occuper avaient la possibilité d'être faites chez moi à domicile.



2. Les tâches accomplies durant la période du stage

2.1 Création du système d'affiliation marketing

Introduction au projet d'affiliation

La première tâche que j'avais à faire était de l'affiliation marketing. Dès le début du stage, j'ai eu une visioconférence sur signal le premier jour où nous avons parlé de tout ce que j'allais faire dans le cadre de l'entreprise, des missions que nous allions aborder et de ce qu'il attendait de moi pendant la période de mon stage. La visioconférence se déroulait à trois avec Christophe LEMEE, Bilal, un autre élève de MMI3 à Gustave Eiffel, et moi-même. J'ai alors eu accès à une boîte mail pour les membres de l'entreprise ou l'on m'a envoyé un fichier pdf avec tous les attendus et les missions que j'avais à faire avec comme consigne principale : Créer un programme d'affiliation complet (mécanisme, rémunération, communication, visuels, référencement Marketplace, ...) pour le service de dépôt de capital, vers les centres d'affaires, les sociétés de domiciliations, les formalistes juridiques et les cabinets d'avocats spécialistes dans la création d'entreprises. La deuxième partie de cette consigne concerne le scrapping qui est une autre tâche que j'ai dû effectuer en parallèle et dont je parlerais plus tard dans ce rapport. Les actions que j'avais à mener pour cette première mission étaient :

- Préparer le programme d'affiliation dans SIO
- Référencer le programme sur la Marketplace SIO (voir avec Bilal)
- Référencer le programme sur la Marketplace <https://www.clickbank.com> ainsi que regarder un tuto youtube clickbank
- Créer un template pour la partie emailing de SIO, afin de pousser l'offre auprès des cibles Google Maps (attention ne pas shooter l'emailing sans validation de Christophe)

L'une des directives essentielles de cette mission était d'intégrer l'utilisation de l'intelligence artificielle (IA) dans mon travail. On m'avait fourni une liste d'outils IA que je devais absolument utiliser pour l'élaboration du programme d'affiliation, mais aussi pour les tâches de scrapping, dont je parlerai plus en détail dans une autre section de ce rapport.

Lors de cette première visioconférence, j'ai également appris que nous allions travailler sur des projets communs en rapport avec le site internet de l'entreprise, en particulier en ce qui concerne l'affiliation marketing. Christophe nous a expliqué les raisons pour lesquelles il souhaitait développer un programme d'affiliation et la manière dont il envisageait de le promouvoir à l'échelle nationale. Le programme d'affiliation devait initialement se concentrer sur le service de dépôt de capital, un nouveau service ajouté aux prestations de l'entreprise.

Le concept du dépôt de capital consiste à placer une somme d'argent dans une entreprise ou un service financier, souvent dans le but de bénéficier d'avantages spécifiques, de générer des intérêts, ou de sécuriser des transactions futures. L'idée était de rendre ce service plus visible et accessible grâce à un réseau d'affiliés, qui seraient rémunérés pour chaque client qu'ils apporteraient à l'entreprise.

Le site web de Deep Block, sur lequel repose une grande partie de l'activité de l'entreprise, est structuré et hébergé grâce à une plateforme appelée Système.io (SIO). Contrairement à WordPress, SIO est une solution intégrée qui permet de gérer non seulement la création de sites web, mais aussi des tunnels de vente, des campagnes d'emailing, et des programmes d'affiliation. Système.io se distingue par sa simplicité d'utilisation et son approche tout-en-un, qui évite le besoin de plugins ou de développements personnalisés souvent nécessaires avec des CMS plus traditionnels comme WordPress.



Je suis arrivé à Chessy, dans l'un des nouveaux bâtiments de la place octogonale, et j'ai été agréablement surpris par le quartier. Il offre de nombreux services : restaurants, commerces, et il est à seulement 5 minutes de Val d'Europe et de la Gare du RER A. Chaque matin, je me levais vers 7h40, me préparais, et partais entre 8h et 8h20 pour arriver à Chessy à 10h en prenant le bus jusqu'à Noisy-le-Grand, puis le RER A jusqu'à Chessy.

Nous étions installés dans de nouveaux locaux de Village by CA. À mon arrivée, j'ai découvert que nous étions quatre sur le site : Hugo, un master en informatique chargé du développement d'une extension web pour les signatures électroniques et la gestion de documents ; Bilal, étudiant à Gustave Eiffel sur le site de Meaux en MMI3, qui travaillait sur l'optimisation du site de l'entreprise, le design et les fonctionnalités ; et Fiona, une étudiante en graphisme et arts numériques, qui gérait la création de visuels pour divers projets. Hugo commençait à 9h, car il avait les clés, tandis que Bilal et moi commencions à 10h.

Dès les premiers jours de mon stage, j'ai commencé à explorer Système.io pour comprendre son fonctionnement, notamment en ce qui concerne la gestion des canaux de vente et les types de produits qui peuvent y être commercialisés. Bilal, qui avait déjà une certaine expérience avec cette plateforme, m'a guidé à travers les différentes fonctionnalités, en particulier celles liées à la création et à la gestion du site web de l'entreprise. Nous avons également visionné plusieurs tutoriels, certains fournis par Christophe et d'autres trouvés sur YouTube, pour approfondir notre connaissance de Système.io et de ses capacités.

Développement du programme d'affiliation

Après environ deux jours, je me sentais suffisamment à l'aise pour commencer à prendre des initiatives. Nous avons alors entrepris de créer le programme d'affiliation centré sur le service de dépôt de capital. L'affiliation marketing repose sur un système dans lequel une entreprise récompense des partenaires, appelés affiliés, pour avoir générée des ventes ou des inscriptions grâce à des liens d'affiliation spécifiques. Ces liens, uniques pour chaque affilié, permettent de suivre leurs efforts de promotion et de les récompenser en conséquence. Pour mettre en place ce programme, j'ai d'abord conçu un template de mail, qui serait automatiquement envoyé à chaque personne souhaitant rejoindre le programme d'affiliation. Ce template a été créé à partir de banques d'images fournies par Fiona, la graphiste de l'entreprise. Chaque mail était personnalisé en fonction du destinataire, avec des liens d'affiliation uniques pour chaque client ou partenaire. J'ai veillé à ce que les images utilisées soient libres de droits et qu'elles correspondent à l'identité visuelle de l'entreprise.

Cependant, j'ai commis une erreur lors de la création de ce mail. Au lieu de mettre l'accent sur le programme d'affiliation, j'ai parmis en avant le service de dépôt de capital lui-même. Cela donnait l'impression que le mail cherchait à vendre le service de dépôt de capital, alors que l'objectif principal était de recruter des affiliés pour promouvoir ce service. Malheureusement, étant déjà engagé dans une autre mission à ce moment-là, je n'ai pas pu corriger cette erreur immédiatement.



Malgré cette erreur, nous avons continué à progresser dans le développement du programme d'affiliation. Cependant, nous avons rapidement rencontré un obstacle majeur : la structure actuelle du site web de l'entreprise ne permettait pas de relier les modules de paiement du dépôt de capital aux liens d'affiliation. Cette limitation technique empêchait de créditer correctement les commissions aux affiliés après chaque vente. Pour surmonter ce

problème, nous avons dû approfondir notre compréhension de la configuration des liens sur Système.io, mais malgré nos efforts, nous n'avons pas réussi à résoudre ce problème avant la fin de mon stage.

Parallèlement à cette tâche, j'ai également remarqué plusieurs défauts sur le site web, tant au niveau esthétique qu'au niveau des textes. J'ai pris l'initiative de corriger ces erreurs et d'apporter des améliorations là où cela était nécessaire. Le lien vers le site web officiel de l'entreprise est le suivant : <https://www.deepblock.eu/>.

Il est également important de noter que nous avions initialement prévu de créer le programme d'affiliation sur ClickBank, une autre marketplace bien connue dans le domaine du marketing d'affiliation. Cependant, en raison des problèmes rencontrés avec Système.io et de la priorité donnée à cette plateforme, nous avons décidé de concentrer nos efforts sur la résolution des problèmes sur Système.io avant de passer à ClickBank. Néanmoins, nous nous étions informés sur la façon dont nous allions procéder sur ClickBank, et nous avions une stratégie en place, prête à être déployée une fois les problèmes sur SIO résolus.

Tout au long de ce processus, nous avons régulièrement fait des compte-rendus à Christophe, qui nous a fourni des directives pour surmonter les obstacles rencontrés et des idées pour améliorer notre façon de travailler. Ses retours nous ont aidés à améliorer notre approche et à rester concentrés sur les objectifs à atteindre, malgré les problèmes rencontrés.

2.2 Scraping des données

Introduction au scraping

Dans la fiche de mission qui m'avait été fournie le premier jour, l'une des missions principales consistait à effectuer du scraping d'emails de certains domaines d'activité spécifiques. Cela devait permettre d'effectuer un shooting (envoi massif d'emails) et de lancer la prospection pour le système d'affiliation lié au dépôt de capital. Pour rappel, l'objectif principal de mon stage est de : Créer un programme d'affiliation complet (incluant le mécanisme, la rémunération, la communication, les visuels, le référencement sur les marketplaces, etc.) pour le service de dépôt de capital. Ce programme doit cibler des centres d'affaires, des sociétés de domiciliation, des formalistes juridiques et des cabinets d'avocats spécialisés dans la création d'entreprises. Parmi les tâches, il était donc crucial de scrapper des données provenant de domaines liés de près à ce service.

La deuxième partie de la consigne spécifie les domaines d'activité à cibler pour ce travail de scraping. Il était précisé de "Générer à partir de Google Maps la sélection des cibles dans toute la France, et à l'aide du module Importfromweb, de réaliser l'export des adresses courriels en prévision de les importer dans SIO." Une vidéo explicative sur l'utilisation d'Importfromweb (un plugin Google Sheets) m'avait également été fournie pour m'assister dans cette tâche.

J'ai sérieusement commencé la mission de scraping dès la deuxième semaine de stage. Pour bien comprendre la nature de cette tâche, j'ai consulté la vidéo qui m'avait été fournie, ainsi que d'autres vidéos pour approfondir ma compréhension du scraping, notamment comment l'opérer efficacement sur Google Maps. Après cela, j'ai installé l'extension Importfromweb, qui est une extension Google Sheets permettant, en théorie, de récupérer les informations de contact et autres caractéristiques d'une entreprise via quelques manipulations spécifiques sur Google Sheets.

Après avoir suivi les tutoriels, j'ai commencé à obtenir des résultats en grand nombre, mais ceux-ci ne correspondaient pas aux attentes. Les informations que je devais récupérer incluaient :

- Le nom de l'entreprise
- L'email de contact
- Le code postal du siège de l'entreprise

Parmi ces informations, l'email était l'élément le plus crucial, suivi du nom de l'entreprise et du code postal. Ces données étaient nécessaires pour mieux cibler les campagnes de shooting, que ce soit exclusivement en Île-de-France ou dans d'autres régions de France. Le principal problème rencontré avec Importfromweb était que cet outil ne parvenait pas à récupérer l'email de contact, qui était pourtant l'information la plus essentielle pour l'envoi massif d'emails.

Après avoir constaté ce problème, j'ai fait un retour à Christophe, mon superviseur, qui m'a alors demandé de tester plusieurs autres sites et extensions qu'il avait identifiés. En parallèle, j'ai également commencé à mener des recherches personnelles pour trouver une méthode plus rapide et efficace pour obtenir les informations demandées.

Parmi les options que j'ai explorées, l'une des plus prometteuses était un service appelé Scrap.io. Ce service semblait répondre parfaitement à nos besoins, mais il avait un inconvénient majeur : il était payant. Pour tenter de contourner cet obstacle, nous avons contacté les responsables de Scrap.io, une entreprise française, en espérant négocier une commande spéciale adaptée à nos besoins. Malheureusement, malgré nos appels répétés avec Christophe, nous n'avons pas reçu de réponse favorable de leur part.

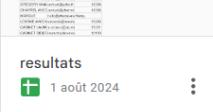
The screenshot shows the Scrap.io homepage with a dark background. At the top, there's a navigation bar with links for Pricing, FAQ, About us, Extension, Help, Login, and Create account. Below the navigation, the main heading is "Google Maps leads generator" with the subtext "Turn Google Maps into a leads machine. Try for free for 7 days and cancel anytime." A search bar at the top has dropdowns for "Restaurant", "Paris", and "France". Below the search bar is a button labeled "Restaurant in Paris (France)". The main content area displays a table of search results:

| Restaurant | Phone | Email | Social Media |
|------------------------------|----------------|-------------------------------|--------------|
| Le Ciel de Paris Restaura... | +33 1*** *7 64 | c***tact@cieldeparis.com | |
| Epicure | +33 1*** *3 40 | r***on@bekeercollection.co... | |
| Le Cinq | +33 1*** *154 | j***par@fourseasons.com | |
| Restaurant Dans le Noir ... | +33 1*** *8 04 | c***noel@danslenoir.com | |
| Benoit Paris | +33 1*** *5 76 | r***mot@ducasse-paris.com | |
| Ellsworth | +33 1*** *9 66 | e***aris@gmail.com | |
| Restaurant Eels | +33 1*** *0 20 | r***eels@gmail.com | |
| Pur' - Jean-François Rou... | +33 1*** *0 60 | r***ants@hvatt.com | |

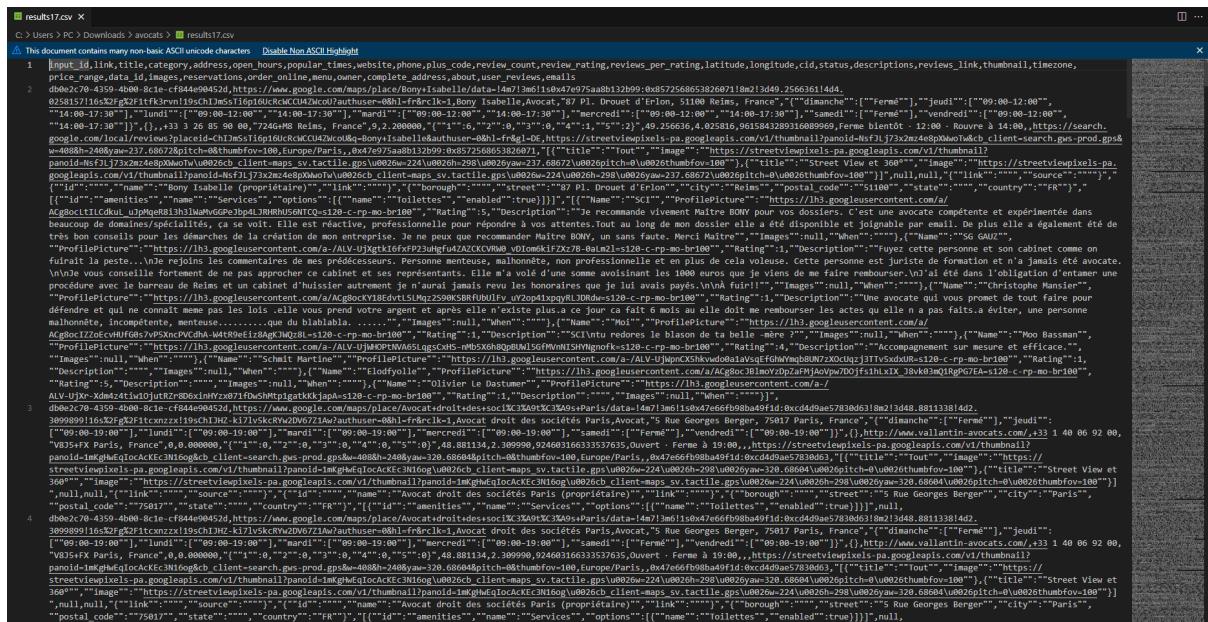
Face à cette impasse, j'ai poursuivi mes recherches pour identifier des alternatives gratuites ou peu coûteuses qui pourraient nous permettre de récupérer les emails de manière efficace et en conformité avec les exigences du projet. Après avoir exploré diverses extensions de scraping, j'ai pris le temps de comparer leurs fonctionnalités et de les évaluer en fonction de leur capacité à extraire précisément les informations dont nous avions besoin. Finalement, j'ai trouvé une extension de scraping qui semblait prometteuse, car elle permettait de récupérer exactement les informations recherchées. Cependant, son utilisation était loin d'être idéale : il fallait naviguer sur Google Maps, sélectionner une zone géographique précise, comme une ville, puis cliquer sur un bouton pour générer un fichier CSV contenant les contacts. Malheureusement, la version gratuite de l'outil ne permettait d'obtenir que 20 contacts par exportation.

En outre, le processus présentait plusieurs inconvénients majeurs. Certaines entreprises n'avaient tout simplement pas d'adresse email disponible, ce qui m'obligeait à trier manuellement et à retirer ces entrées du fichier final. De plus, la nécessité de sélectionner manuellement la zone géographique rendait le processus imprécis, ce qui entraînait la présence de doublons et une couverture incomplète des contacts potentiels. Malgré mes efforts pour rassembler un nombre significatif de contacts en utilisant cette méthode, les résultats étaient décevants : le processus était lent, fastidieux et inefficace pour atteindre l'objectif fixé.

Je ne vais pas approfondir mais comme visible ci-dessus j'ai testé un nombre incalculable de scrappeurs sans grande réussite :

| | | | | |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------|
|  resultats + 1 août 2024 ⋮ |  Feuille de calcul sans titre + 1 août 2024 ⋮ |  Avocats - Feuille 1 + 18 juil. 2024 ⋮ |  Avocats - Feuille 1 - Avoc... + 18 juil. 2024 ⋮ |  Avocats + 17 juil. 2024 ⋮ |
|  Formalistes Juridiques + 15 juil. 2024 ⋮ |  Centres D'affaires + 15 juil. 2024 ⋮ |  Feuille de calcul sans titre + 12 juil. 2024 ⋮ |  export_domiciliation + 11 juil. 2024 ⋮ |  Centres d'affaire + 27 juin 2024 ⋮ |
|  Feuille de calcul sans titre + 27 juin 2024 ⋮ |  Centre d'affaires 5 + 27 juin 2024 ⋮ |  Centre d'affaires 4 + 27 juin 2024 ⋮ |  Centre d'affaires 3 + 27 juin 2024 ⋮ |  Centre d'affaires 2 + 27 juin 2024 ⋮ |
| | | | | |

La situation a changé lorsque nous avons découvert un script disponible sur GitHub, capable d'extraire les informations souhaitées de manière beaucoup plus fiable. Cependant, cette nouvelle méthode posait un autre problème : l'organisation des fichiers générés. Lors de ma première extraction à l'aide de ce script, je me suis retrouvé avec un fichier compressé en .zip contenant 99 fichiers, chacun variant en taille de 0 kilo-octet à 1 Mo. Plus un fichier est volumineux, plus il contient de contacts d'entreprises. Ces 99 fichiers étaient censés être classés par département, mais leur organisation n'était pas aussi simple ni aussi ordonnée qu'escompté.



Voici à quoi ressemblait un fichier CSV typique généré par cette extraction : chaque fichier contenait une grande quantité d'informations, mais dans un format désordonné et difficile à manipuler. Par exemple, le nom de chaque entreprise était associé à un champ nommé "name". Toutefois, un autre champ portant le même nom pouvait contenir le nom du propriétaire de l'entreprise, rendant la distinction entre les deux difficile. Un autre champ, nommé "code_postal", contenait parfois le code postal de l'entreprise, mais cette information n'était disponible que dans environ 90 % des cas. Quant aux adresses email, elles étaient particulièrement difficiles à extraire : elles apparaissent en fin de bloc de texte, sans balise ou identifiant spécifique pour les isoler facilement.

Face à ce désordre, j'ai dû passer une journée entière à extraire manuellement chaque information, utilisant des recherches par mots-clés grâce à la commande "Control + F" pour trouver les éléments nécessaires. Cette tâche était extrêmement fastidieuse, répétitive et exigeante. J'ai dû parcourir des milliers de lignes de texte, espérant trouver un nombre suffisant d'entreprises avec une adresse email valide. Pourtant, même dans les fichiers les plus volumineux, seuls 30 % des entreprises disposaient d'une adresse email sur leur profil Google, ce qui réduisait considérablement l'efficacité de l'extraction. De plus, il n'était pas rare de rencontrer des erreurs de caractères, avec des noms d'entreprises affichés de manière incorrecte, par exemple sous la forme "@lta/oa".

J'ai donc dû corriger manuellement toutes ces erreurs et rassembler les informations valides dans un fichier CSV, en séparant les données par des virgules sur Google Sheets. J'ai choisi d'utiliser Sheets car c'est l'outil que je maîtrise le mieux. L'objectif était de produire un fichier final en format CSV, contenant uniquement les informations pertinentes, car Système.io avait besoin de ce format précis pour importer les contacts de l'entreprise en vue du shooting final. Le fichier final devait contenir le nom de l'entreprise en majuscule, une exigence qui m'avait été spécifiée dès le premier jour de scraping. Pour faciliter ce processus, j'ai utilisé un convertisseur automatique pour transformer les noms en majuscule, plutôt que de le faire manuellement pour chaque entrée.

| | A | B | C |
|----|---------------------------------------------------------------------|-------------------------------------|-------|
| 1 | DOM'A MEAUX | dom-a-meaux@orange.fr | 77100 |
| 2 | CENTRE D'AFFAIRES REIMS BEZANNE | contact@centre-affaires-bezannes.fr | 51430 |
| 3 | DOMICILIATION LYON / LES TRICOLORES | contact@lestricolores.com | 69003 |
| 4 | LE CUBE BUSINESS | contact@lecube-business.com | 6130 |
| 5 | HAUTES-ALPES DOMICILIATION GAP | hautes-alpesdomiciliation@live.fr | 5000 |
| 6 | ARENAS PARTNERS, DOMICILIATION COMMERCIALE NICE, LOCATION DE BUREAU | message@arenas-partners.com | 6200 |
| 7 | DOMICILIATION ENTREPRISE | naandco@orange.fr | 6000 |
| 8 | SELFBURO | contact@selfburo.fr | 6600 |
| 9 | ZGS&DOM | contact@zgs-domiciliation.fr | 6300 |
| 10 | SAP 06 | sap06@outlook.fr | 6000 |
| 11 | ASSIST'ENTREPRISES | contact@assistentreprises.com | 6400 |
| 12 | NOVAFFAIRES | contact@novaffaires.fr | 6300 |
| 13 | PARADIGM | contact@pactams-paradigm.com | 6000 |
| 14 | OPTION BUREAU | contact@option-bureau.com | 6130 |
| 15 | PAGES | contact@domiciliation-2607.com | 26200 |
| 16 | CENTRE D'AFFAIRES ET DE SERVICES VALENTINOIS | casv26@outlook.fr | 26000 |
| 17 | CENTRE D'AFFAIRES A A Z | contact@lyoncentredaffaires.com | 69003 |
| 18 | ESPACE FAURIEL | contact@espacefauriel.fr | 42100 |
| 19 | ELD EN LIGNE DIRECTE CHARLEVILLE-MEZIERES | contact@enlignedirecte.fr | 8000 |
| 20 | DACTYFIL | contact@dactyfil.fr | 10000 |
| 21 | CENTRE D'AFFAIRES - DOMICILIATION - CENTRE DE FORMATION PERPIGNAN | contact@centre-affaires-mg.com | 66330 |
| 22 | CARCA BOX | carcabox@gmail.com | 11000 |
| 23 | BAQAN CENTRE D'AFFAIRES | info@dom-marseille.com | 13008 |
| 24 | LATELIER - DOMICILIATION D'ENTREPRISE | contact@latelier-aix.com | 13100 |
| 25 | DOMICILIATION MARSEILLE 7ÈME | info@domiciliationmarseille7eme.fr | 13007 |
| 26 | DOMICILIATION MARSEILLE 5ÈME | info@domiciliationmarseille5eme.fr | 13005 |

Cependant, le travail n'était pas terminé. Après avoir extrait et corrigé les données pour les premiers fichiers, je les ai soumis à Christophe, qui m'a ensuite fourni deux autres fichiers à traiter. J'ai appliqué la même méthode manuelle pour ces nouveaux fichiers, passant encore des heures à sélectionner, corriger et organiser les données. Mais Christophe a rapidement relevé plusieurs problèmes dans ma première extraction : les caractères spéciaux, les accents et les codes postaux à quatre chiffres posaient problème lors de l'importation des données dans SIO. Par exemple, les codes postaux à quatre chiffres perdaient le zéro initial dans les tableurs, ce qui faussait les données.

Pour résoudre ce problème, j'ai tenté d'ajouter un caractère invisible au début des codes postaux à quatre chiffres, mais cela n'a pas fonctionné comme prévu. C'est à ce moment-là que mon objectif est devenu clair : je devais trouver un moyen d'extraire les fichiers aussi rapidement et proprement que possible, tout en évitant les erreurs et en respectant les contraintes de formatage.

Développement et utilisation d'un script d'extraction de données

Face à ces défis, Christophe m'a suggéré de développer un script capable d'automatiser non seulement l'extraction des données, mais aussi leur formatage : élimination des caractères spéciaux, des accents, gestion correcte des codes postaux, suppression des doublons, et filtrage pour ne conserver que les entreprises avec une adresse email. Le but final était d'obtenir un fichier CSV parfaitement formaté, prêt à être importé en un seul clic dans Systeme.io. Cependant, mes compétences en Python étaient limitées, et nous n'avions pas eu suffisamment de cours en MMI pour créer un script de cette complexité.

J'ai donc commencé à me documenter intensivement, cherchant des ressources et des exemples de code pouvant m'aider à accomplir cette tâche. J'ai également sollicité l'aide d'Hugo, un développeur expérimenté qui travaillait avec moi sur le projet. Grâce à nos efforts conjoints, la première version du script a pu voir le jour. Cette version permettait d'extraire toutes les données et de les placer dans un fichier CSV, en utilisant les bibliothèques Python "csv" et "zipfile". Cependant, cette version initiale n'était pas sans défauts : les doublons persistaient, les problèmes de caractères et de codes postaux n'étaient pas résolus, et une grande partie du travail de nettoyage des données devait encore être effectuée manuellement.

Malgré ces limitations, cette première version représentait une avancée considérable par rapport à la méthode manuelle précédente. Là où il me fallait 10 heures pour traiter les données, je pouvais maintenant le faire en une heure, bien que le but ultime reste de permettre un export en un clic, ce qui était encore loin d'être atteint. Avec cette première version du script en place, nous avions une bonne base pour poursuivre le développement et améliorer notre méthode d'extraction des données. Mon objectif était de perfectionner le script pour qu'il puisse gérer automatiquement les exceptions et les erreurs que nous avions identifiées, en éliminant les étapes manuelles restantes.

Le processus de développement du script a été itératif et a nécessité de nombreuses modifications pour atteindre l'efficacité recherchée. Nous avons intégré des fonctions pour détecter et supprimer les doublons, gérer correctement les codes postaux à quatre chiffres en les supprimant simplement sans autre possibilité, et nettoyer les noms d'entreprises en supprimant les caractères spéciaux et les en remplaçant les accents. Au fil des itérations, le script est devenu de plus en plus utilisable, capable de produire un fichier CSV propre et bien formaté, prêt à être importé sans autre intervention manuelle. Voici à quoi ressemble le script final :

```
import re
import csv
import os
import zipfile
import tarfile
```

```

def remove_accents(input_str):
    """
    Retire les accents d'une chaîne de caractères en les remplaçant par leurs équivalents non accentués.
    """

    accents = {
        'a': 'àáâãäåÀÁÂÃÄÅ',
        'e': 'èéêëÈÉÊË',
        'i': 'íîñõíïÌÍÑÕ',
        'o': 'òóôõöÒÓÔÕÖ',
        'u': 'ùúûüÙÚÛÜ',
        'y': 'ýÿÝ'
    }
    for char, accented_chars in accents.items():
        for accented_char in accented_chars:
            input_str = input_str.replace(accented_char, char)
    return input_str

# Fonction pour décompresser les fichiers ZIP
def unzip_file(zip_filepath, extract_to):
    with zipfile.ZipFile(zip_filepath, 'r') as zip_ref:
        zip_ref.extractall(extract_to)

# Fonction pour décompresser les fichiers TGZ
def untar_file(tar_filepath, extract_to):
    with tarfile.open(tar_filepath, 'r:gz') as tar_ref:
        tar_ref.extractall(extract_to)

# Déterminer le type de fichier et décompresser
def extract_files(file_path, extract_to):
    if file_path.endswith(".zip"):
        unzip_file(file_path, extract_to)
    elif file_path.endswith(".tgz") or file_path.endswith(".tar.gz"):
        untar_file(file_path, extract_to)
    else:
        raise ValueError("Unsupported file type")

# Demander à l'utilisateur de sélectionner le fichier compressé
file_path = input("Entrez le chemin du fichier ZIP/TGZ/RAR à traiter : ")
extract_to = "./extracted_files"

# Décompresser le fichier

```

```

extract_files(file_path, extract_to)

str_outputfrance = ""
str_outputrp = ""

# Parcourir les fichiers extraits
files = os.listdir(extract_to)

# Obtenez le nom du dossier en cours
current_dir = os.path.basename(os.getcwd())

for file in files:
    file_path = os.path.join(extract_to, file)
    if file.endswith(".csv"):
        with open(file_path, newline="", encoding='utf-8') as csvfile:
            spamreader = csv.reader(csvfile, delimiter=',', quotechar='|')
            for row in spamreader:
                if len(row) > 5 and row[-1] != "" and row[-1] != "emails":
                    codepostal = row[5][:6]
                    nom = remove_accents(row[2]).upper()
                    email = row[-1]
                    if all(term not in email for term in ["domain", "wixpress", ".png", "ovh"]):
                        if "{" not in nom and "%" not in nom:
                            if "{" not in codepostal and len(codepostal) >= 5:
                                email = email.replace("", " ")
                                nom = re.sub(r'[+\-\.\"]', " ", nom)
                                codepostal = codepostal.replace('+', "").replace("", " ")
                                if any(cp in codepostal for cp in ["75", "77", "78", "91", "92", "93", "94", "95"]):
                                    str_outputrp += f"{nom},{email},{codepostal}\n"
                                else:
                                    str_outputfrance += f"{nom},{email},{codepostal}\n"

output_file_france = f'{current_dir}_france.csv'
with open(output_file_france, 'w', newline="", encoding='utf-8') as csvfile:
    csvfile.write("Nom,Email,Code Postal\n")
    csvfile.write(str_outputfrance)

output_file_rp = f'{current_dir}_rp.csv'
with open(output_file_rp, 'w', newline="", encoding='utf-8') as csvfile:
    csvfile.write("Nom,Email,Code Postal\n")
    csvfile.write(str_outputrp)

```

```
print(f"Résultats enregistrés dans '{output_file_france}' et '{output_file_rp}'")
```

Pour rentrer dans les détails du code, ce script Python que j'ai développé a pour but de simplifier et d'automatiser le traitement de fichiers, il est découpé en plusieurs parties :

1. **Suppression des accents** : Une première étape consiste à nettoyer les noms des entreprises en enlevant tous les accents (par exemple, "é" devient "e"). Cela permet de normaliser les données et d'éviter des erreurs lors de l'importation dans d'autres systèmes.
2. **Décompression des fichiers** : Le script est capable de prendre en charge des fichiers compressés au format ZIP ou TGZ. Ces fichiers contiennent généralement plusieurs fichiers CSV (un format de fichier qui stocke des données tabulaires, comme une feuille Excel). Le script extrait automatiquement ces fichiers compressés dans un dossier spécifique.
3. **Lecture et tri des données** : Une fois les fichiers décompressés, le script examine chaque fichier CSV pour en extraire les informations importantes : le nom de l'entreprise, son adresse email, et son code postal. Il filtre également les informations pour supprimer les doublons et s'assurer que seules les données pertinentes sont conservées.
4. **Classification géographique** : Ensuite, le script classe les entreprises en deux catégories : celles qui se trouvent en région parisienne (selon leur code postal) et celles qui se trouvent dans le reste de la France.
5. **Création des fichiers finaux** : Enfin, les informations triées sont enregistrées dans deux nouveaux fichiers CSV, prêts à être utilisés pour d'autres opérations. L'un des fichiers contient les entreprises situées en région parisienne, et l'autre celles situées dans le reste de la France.

Après des semaines de travail acharné, j'ai finalement atteint mon objectif : créer un script capable d'extraire et de traiter les données en un seul clic. Ce script, bien qu'il m'ait demandé plusieurs heures de développement et de tests, s'est révélé être une solution efficace pour automatiser une tâche qui, auparavant, me prenait un temps considérable et était extrêmement fastidieuse à réaliser manuellement.

Au début du projet, je n'avais pas envisagé cette solution, ce qui m'a conduit à passer beaucoup de temps à extraire et nettoyer les données de manière manuelle. Avec le recul, je regrette de ne pas avoir eu cette idée plus tôt, car elle m'aurait permis de gagner énormément de temps et d'efforts. Cependant, la bonne nouvelle est que le script est désormais terminé et opérationnel. Il est non seulement fonctionnel pour les besoins actuels, mais il pourra être utilisé à l'avenir pour des projets de l'entreprise.

Le script fonctionne à 99 %, ce qui signifie qu'il est extrêmement fiable. Toutefois, il reste quelques petites imperfections. Lors de la comparaison entre les résultats obtenus manuellement et ceux générés par le script, j'ai remarqué une légère perte d'information et quelques contacts qui n'ont pas été capturés. Ces pertes sont minimes, et les avantages en termes de gain de temps et de réduction des erreurs humaines compensent largement ces

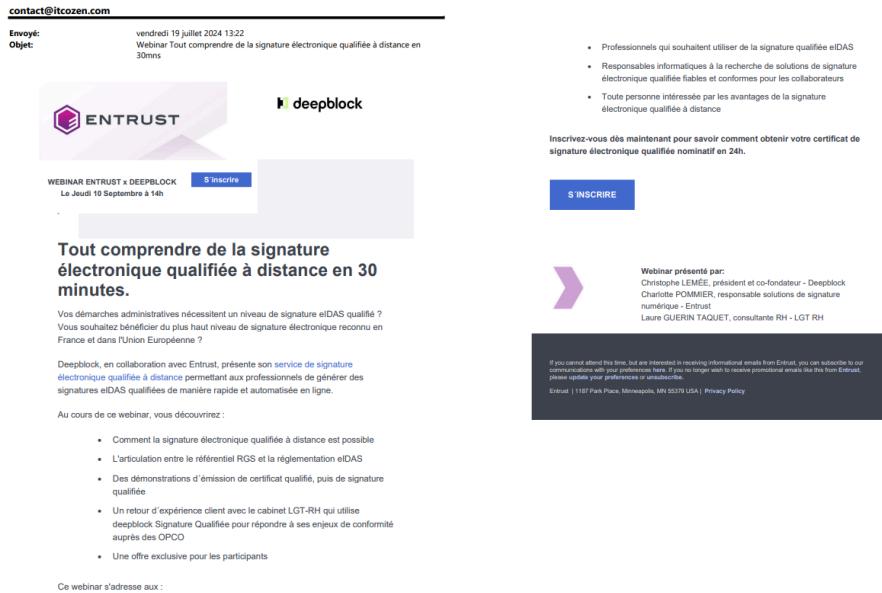
petites omissions. Lorsque j'ai présenté le script à Christophe, il a été très satisfait du résultat. J'ai pu lui fournir les dernières extractions nécessaires peu de temps après avoir finalisé le script. Cela a permis d'ajouter rapidement et efficacement tous les contacts dans le système de gestion des contacts (SIO), ce qui a marqué la fin de cette mission.

En résumé, cette mission m'a non seulement permis d'améliorer mes compétences en programmation, mais il a également démontré l'importance de l'automatisation dans les processus de travail. Le script que j'ai développé est une solution qui continuera à apporter de la valeur à l'entreprise en simplifiant et en accélérant des tâches similaires à l'avenir.

2.3 Conception de l'affiche pour le séminaire en ligne

Contexte et objectifs du projet

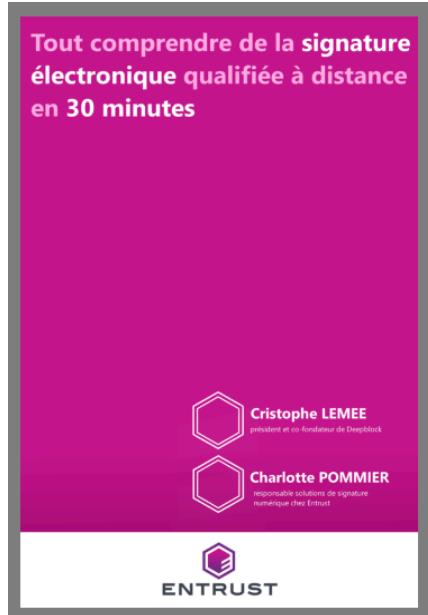
Vers la fin de mon stage, j'ai eu l'occasion de travailler sur un projet créatif en réalisant une affiche. Christophe m'a confié plusieurs fichiers PDF contenant toutes les informations nécessaires concernant un séminaire en ligne, organisé en partenariat entre Deep Block et Entrust. Mon rôle était de concevoir une affiche en respectant scrupuleusement la charte graphique d'Entrust. C'était un projet important, car l'affiche devait non seulement capter l'attention, mais aussi refléter fidèlement l'image des deux entreprises.



Création de la première version de l'affiche

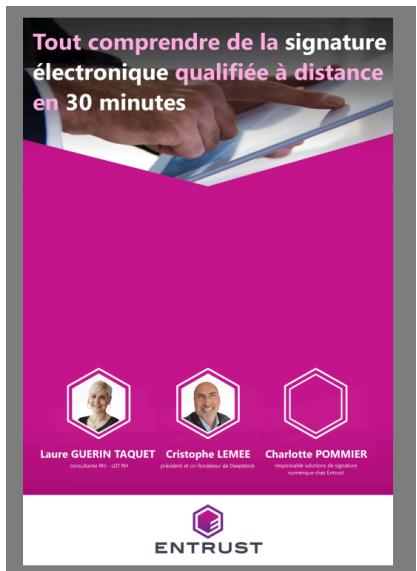
Pour commencer, j'ai étudié les couleurs principales de la marque Entrust afin de les intégrer dans ma création (grâce à une charte graphique qu'il m'avait fourni). J'ai également effectué des recherches approfondies sur l'entreprise ainsi que sur les personnes qui devaient intervenir en tant qu'orateurs lors du séminaire. Avant de me lancer dans la conception, je me suis familiarisé avec les principes de base des séminaires en ligne et les éléments clés que l'on met généralement en avant sur les affiches promotionnelles de ces

événements. En m'inspirant de modèles trouvés sur Internet, j'ai décidé de créer un fond dégradé aux couleurs d'Entrust pour établir une base visuelle solide. J'ai conçu une première version de l'affiche en y intégrant un titre accrocheur, une image représentative et des espaces réservés pour les photos des intervenants. Une fois cette première version terminée, je l'ai présentée à Christophe, qui a immédiatement approuvé la direction artistique que j'avais choisie.



Révisions, ajustements et finalisation de l'affiche

Cependant, au fur et à mesure que le projet avançait, Christophe m'a informé d'un autre intervenant du séminaire, ce qui nécessitait de revoir la mise en page initiale. J'ai donc ajusté le plan de l'affiche et demandé les photos des participants pour pouvoir les intégrer de manière cohérente. Après plusieurs révisions, j'ai envoyé à Christophe une version intermédiaire de l'affiche, dans laquelle j'avais placé une image libre de droits, trouvée sur une banque d'images en ligne, en forme de "V" en haut de la page pour structurer visuellement le contenu. Ce design, visible dans l'ébauche ci-dessous, a reçu l'approbation de Christophe, qui m'a encouragé à poursuivre dans cette direction.



À ce stade, j'ai finalisé l'affiche en y ajoutant toutes les informations importantes concernant le séminaire et en intégrant la dernière photo du nouvel intervenant que Christophe m'avait envoyée. Le processus de création m'a pris environ trois heures, mais c'était une tâche que j'ai particulièrement appréciée. J'ai trouvé ce travail extrêmement gratifiant, car il combine des compétences techniques et créatives. Après de nombreuses discussions et ajustements en temps réel, j'ai pu livrer une affiche finale qui répondait pleinement aux attentes de Christophe. Ce projet reste l'un de mes favoris parmi ceux que j'ai réalisés au cours de mon stage.



Conclusion

Mon stage au sein de l'entreprise Deepblock m'a offert une opportunité précieuse de mettre en pratique les connaissances théoriques acquises au cours de ma formation en Bachelor MMI. En travaillant dans une entreprise innovante spécialisée dans l'intelligence artificielle et la blockchain, j'ai pu développer des compétences dans le domaine du marketing digital, notamment à travers la création d'un programme d'affiliation marketing et l'utilisation d'outils d'intelligence artificielle.

La réalisation de ces projets m'a permis de renforcer mes compétences en gestion de projets numériques, en communication, ainsi qu'en résolution de problèmes techniques, particulièrement face aux défis liés à la plateforme Système.io et au scraping. J'ai particulièrement apprécié le développement du script Python pour le scraping, qui a été une occasion stimulante d'appliquer mes compétences en programmation à des tâches concrètes et complexes.

De plus, l'expérience du télétravail et du travail en présentiel m'a aidé à développer une meilleure autonomie et une capacité d'adaptation aux environnements de travail. Ce stage a confirmé mon intérêt pour le domaine du marketing digital et m'a donné un aperçu concret de ce qu'il se passait dans les start-ups technologiques où j'aimerais travailler plus tard.

Glossaire

- **Affiliation marketing** : Technique de marketing où une entreprise rémunère des partenaires (affiliés) pour générer des ventes ou des inscriptions via des liens d'affiliation spécifiques.
- **Système.io (SIO)** : Plateforme tout-en-un pour la création de sites web, tunnels de vente, campagnes d'emailing, et gestion de programmes d'affiliation.
- **ClickBank** : Marketplace spécialisée dans les produits numériques et les programmes d'affiliation, permettant aux entreprises de proposer leurs offres et aux affiliés de promouvoir ces offres.
- **Scraping** : Technique utilisée pour extraire des données de sites web automatiquement à l'aide de scripts ou d'outils spécialisés.
- **Importfromweb** : Extension Google Sheets permettant d'extraire des informations de contact de sites web tels que Google Maps.
- **Scrap.io** : Service de scraping payant permettant d'extraire des données de manière automatisée.
- **Python** : Langage de programmation populaire, utilisé pour le développement de scripts, l'analyse de données et plus encore.
- **CSV (Comma-Separated Values)** : Format de fichier permettant de stocker des données tabulaires sous forme de texte, avec des valeurs séparées par des virgules.
- **ZIP** : Format de fichier compressé permettant de regrouper plusieurs fichiers dans un seul fichier pour faciliter le stockage et le transfert.
- **TGZ (Tarball GZ)** : Format de fichier compressé utilisant l'archivage Tar et la compression Gzip, souvent utilisé pour regrouper et compresser des fichiers.
- **Extension Google Sheets** : Outil ou fonctionnalité ajoutée à Google Sheets pour améliorer ses capacités, comme l'extraction de données depuis le web.

- **Regex (Regular Expressions)** : Langage de programmation pour rechercher et manipuler des chaînes de caractères en utilisant des motifs.
- **Données tabulaires** : Données organisées en lignes et colonnes, souvent utilisées dans des fichiers CSV ou des bases de données.
- **Décompresser** : Processus de restauration de fichiers compressés à leur état d'origine pour permettre leur utilisation.
- **API (Application Programming Interface)** : Interface qui permet à différents logiciels de communiquer entre eux en exposant des fonctionnalités ou des données.
- **Script** : Programme ou ensemble d'instructions écrites pour automatiser des tâches spécifiques, souvent dans un langage de programmation comme Python.