



ÉTHIQUE ET ASPECTS LÉGAUX

RAPPORT
-
STRITCRID

Plateforme de recrutement basée sur l'IA

Élèves :

AKOUMBA Ludivine
BOGALE Tegest
FERCHICHI Farouk
LOGAN Victoria
MELLY Jérémie

Enseignant :

PEÑA Carlos

25 janvier 2023

Table des matières

1	Introduction	2
2	Service étudié	3
2.1	Description du sujet	3
2.2	Justification du choix du sujet	3
2.3	Fonctionnement et utilisation du service	3
3	Aspects légaux	4
3.1	Introduction à la PI	4
3.2	Les brevets	5
3.3	Protection des logiciels, autres stratégies de protection	5
3.4	Liberté d'exploitation - Titularité	7
3.5	Contrats et litiges en matière de Propriété intellectuelle	7
3.6	Protection des données	8
4	Aspects éthiques	11
4.1	Biais du machine learning	11
4.2	Position du/de la futur.e employé.e	11
5	Conclusion	12

1 Introduction

Dans le cadre du cours *Éthique et Aspects Légaux (EAL)*, il nous a été demandé d'analyser les aspects éthiques et légaux liés au développement d'un produit ou d'un service et de proposer des considérations et des décisions basées sur les différents aspects étudiés en classe, tels que le design, la production, ou encore la commercialisation.

Nous sommes une équipe de cinq futur.e.s ingénieur.e.s en informatique en cinquième année de Bachelor avec différentes spécialisations :

- 2 étudiant.e.s en ingénierie logicielle ;
- 3 étudiant.e.s en ingénierie des données.

Il nous a également été demandé de nommer une personne responsable (porte-parole), ainsi qu'un.e remplaçant.e dans le cas où le/la responsable serait indisponible. Nous avons donc nommé Victoria Logan comme responsable et Ludivine Akoumba comme remplaçante.

Le sujet que nous avons donc choisi d'investiguer est une **plateforme de recrutement basée sur l'intelligence artificielle (IA)**, comme décrit en détail dans la section suivante (section 2).

2 Service étudié

2.1 Description du sujet

Comme vu par certain.e.s d'entre nous en cours de Gouvernance des Données (GOD) le semestre passé (BA4), il existe des plateformes permettant l'aide au recrutement d'employé.e.s pour les entreprises. Ces plateformes se basent sur un modèle d'apprentissage automatique (machine learning) pour déterminer quel.le candidat.e il serait optimal d'engager. Nous allons donc ici nous intéresser à ce sujet en termes d'aspects éthiques et légaux en les analysant et en proposant des considérations et décisions basées sur ces aspects.

2.2 Justification du choix du sujet

Ce sujet nous intéresse particulièrement car la technologie et ses utilisations nous semblent se diriger vers cette direction ; en tant que futur.e employé.e, on devra probablement - selon nous - y faire face. De plus, étant en études dans le domaine de l'informatique, nous sommes conscient.e.s des failles et des limites de la digitalisation, comme les biais dans les algorithmes et dans les données qui peuvent fausser les résultats. Nous trouvons ainsi qu'il est pertinent d'approfondir le sujet pour en connaître d'avantage.

2.3 Fonctionnement et utilisation du service

Le service que nous avons imaginé étudier ici se base sur les services existants que nous avons trouvés lors de nos recherches en ligne, tels que Avrio AI, Manatal, ou encore Workable. Voici ainsi comment nous l'avons imaginé fonctionner :

- Chaque CV reçu est stocké dans une même base de donnée ;
- Dans une première étape d'aide au recrutement, une analyse de CV est faite via un modèle de machine learning préentraîné. Ce modèle fournit des scores (note sur 10) pour les catégories qu'il a repérées, telles que les expériences du.de la candidat.e, pouvant ainsi par exemple donner un score de "management", ou encore de "années d'expérience dans le domaine". Pour donner un autre exemple de score fourni par le modèle, il classera également le niveau dans certain domaine, par exemple son niveau en C++, en Java, et autres langages de programmation. Ceci sera fait via le machine learning, qui fournira ces score suivant les phrases et leur formulation (contexte) trouvées dans le CV du.de la candidat.e ;
- Dans une deuxième étape d'aide au recrutement, on utilisera un modèle de machine learning permettant de fournir un score de comparaison entre le profil d'un CV établi précédemment et le profil recherché pour le poste en question.

C'est sur ce fonctionnement général que nous allons nous baser pour analyser les aspects légaux (section 3 de ce rapport) et les aspects éthiques (section 4 de ce rapport) inhérents.

Finalement, il est important de préciser que nous avons imaginé cette application sous forme de service Suisse, c'est-à-dire que l'on simule ici le fait d'être une entreprise qui aimerait mettre ce service en place pour l'utilisation d'entreprises tierces Suisses.

3 Aspects légaux

3.1 Introduction à la PI

Tout d'abord, avant de trancher quel type de propriété intellectuelle (PI) est le mieux adapté pour notre projet, nous souhaitons donner une brève définition de la PI ainsi que présenter les différentes possibilités qui sont à dispositions et qui assurent la protection de notre projet.

Le rôle principal de la propriété intellectuelle est la protection d'une nouvelle invention ; l'obtention d'un titre de PI donne l'exclusivité sur l'exploitation des nouvelles découvertes et le droit d'empêcher les autres d'exploiter ces découvertes - ce qui permet en outre de limiter la concurrence dans son domaine, amortir sa RD, augmenter ses marges ainsi que réduire les risques face aux brevets de concurrents.

Il existe plusieurs types de titre de PI, parmi lesquels on retrouve principalement :

Les brevets : Il s'agit d'une forme de droit accordé par le gouvernement à un inventeur ou à son ayant cause, donnant au propriétaire le droit d'interdire à d'autres de fabriquer, d'utiliser, de vendre, d'offrir de vendre et d'importer une invention pendant une période limitée de temps, en échange de la divulgation publique de l'invention. Une invention est une solution à un problème technologique spécifique, qui peut être un produit ou un procédé, et doit généralement répondre à trois exigences principales : elle doit être nouvelle, non évidente et avoir une applicabilité industrielle. Afin d'enrichir l'ensemble des connaissances et de stimuler l'innovation, les titulaires de brevets ont l'obligation de divulguer au public des informations précieuses sur leurs inventions ;

Les droits d'auteurs : Ce sont eux qui confèrent au créateur d'une œuvre originale des droits exclusifs sur celle-ci, généralement pour une durée limitée. Le droit d'auteur peut s'appliquer à un large éventail de formes créatives, intellectuelles ou artistiques, ou œuvres. Le droit d'auteur ne couvre pas les idées et les informations elles-mêmes, seulement la forme ou la manière dont elles sont exprimées ;

Les marques : Une marque est un signe, un dessin ou une expression reconnaissable qui distingue les produits ou services d'un commerçant particulier des produits ou services similaires d'autres commerçants.

Il est important de noter que la liste des PIs ci-dessus est loin d'être exhaustive ; il existe d'autres titres tels que le désigne, ou encore le trade secret (secret commercial). On a cependant ici décidé d'exposer ceux qu'il peut être intéressant d'investiguer pour notre projet.

3.2 Les brevets

De sa définition informatique, une plateforme désigne un matériel ou un logiciel qui héberge une application ou un service. De ce fait, elle fait partie des exclusions de brevetabilité. De plus, pour qu'un service puisse être brevetable, il doit remplir certains critères de brevetabilité :

- **L'utilisabilité industrielle** : ce critère est rempli par notre service car il pourra être utilisé par des entreprises différentes et pourra aider au recrutement de profils de différents corps de métier.
Par contre, il ne respecte pas les critères de :
- **Nouveauté** : car il existe déjà - sur le marché Suisse et mondial - de nombreuses plateformes de recrutement basée sur l'IA.
- **Activité inventive** : Aucune activité inventive ne résulte de l'usage de cette plateforme.

Seulement 1 critère sur 3 de brevetabilité étant respecté, il n'est donc pas possible de breveter notre service.

3.3 Protection des logiciels, autres stratégies de protection

Notre produit/service étant constitué majoritairement de software, la protection par le droit d'auteur.e entre en vigueur dès la création de notre service. Il n'y a pas de formalités à respecter pour obtenir cette protection. En revanche la protection par le droit d'auteur.e est limitée dans le temps. En Suisse, elle est valide jusqu'à 50 ans après la mort de l'auteur.e. Le droit d'auteur.e n'offre pas une protection absolue contre la contrefaçon mais elle nous donne des outils légaux pour nous défendre en cas de litige.

Il est fort probable que, pendant l'implémentation de notre service, nous utilisions des dépendances externes comme :

- React
- PyTorch
- Tensorflow
- etc...

Ces bibliothèques/frameworks sont disponible en libre service sur internet pour une utilisation dans le cadre privé. Les choses se compliquent lorsque l'on veut utiliser de tels outils à des fins commerciales. En effet, même si ces dépendances sont librement accessible par tout.e un.e chacun.e, elles sont quand même protégées par des licences Open Source donc voici quelques exemples :

- **GNU (GPL)** : N'importe quel Software qui utilise une dépendance protégée par cette licence doit être publié en Open Source. Ce qui veut dire que l'on doit publier notre code source et céder tous nos droits à la modification et à la distribution de notre code. Cela ne convient pas à un cadre commercial.

- **Apache** : Avec cette license, nous pouvons librement disposer du code. La seule contrainte est de respecter les conditions de la license Apache : <https://www.apache.org/licenses/LICENSE-2.0>
- **MIT License** : Une des licences les plus permissives. Nous pouvons disposer librement du code tant que l'on ajoute une copie de la license originale MIT et un copyright. Mais certains développeurs l'évitent car elle ne spécifie pas explicitement les droits de distribution.

Nous devons donc apporter une attention particulière aux licences de nos dépendances lors de la conception de notre service.

Comme dit précédemment (3.2 Les brevets), notre plateforme, en elle-même, ne remplit pas les critères pour le dépôt d'un brevet. En revanche, elle sera composée de plusieurs éléments qui, individuellement, pourraient être protégés par les trade secrets.

Comme par exemple :

- Les algorithmes utilisés dans notre modèle de Machine Learning
- Les méthodes de prétraitement des données
- Les données d'entraînement de notre modèle

Ces éléments peuvent être considéré comme des secrets commerciaux et donc peuvent être protégés par le droit des secrets d'affaires.

Afin d'assurer la protection de ces trade secrets nous pouvons :

- **Mettre en place des politiques de confidentialité et de sécurité** : Mettre en place des politiques de confidentialité et de sécurité pour protéger les informations confidentielles et garantir que seuls les employés et les partenaires autorisés ont accès aux informations confidentielles.
- **Utiliser des accords de confidentialité** : (Non Disclosure Agreement) Faire signer des accords de confidentialité à tous les employés, partenaires et prestataires de services pour s'assurer qu'ils comprennent leurs obligations en matière de confidentialité et de protection des informations confidentielles et se protéger légalement.
- **Utiliser des technologies de sécurité** : Mettre en place des technologies de sécurité pour protéger les informations confidentielles, comme le chiffrement des données et l'authentification des utilisateurs.

Si nous trouvons un algorithme de machine learning innovant, une publication défensive pourraient être envisagée. Tout dépendra de l'état du marché au moment de la découverte d'un tel algorithme. En effet, une publication défensive va mettre notre invention dans le domaine public ce qui a plusieurs avantages :

- Se protéger contre les autres brevets de tiers.
- Rendre les caractéristiques de l'invention disponible pour l'industrie (ce qui nous importe peu dans notre cas, nous n'essayons pas définir un standard).
- Créer une preuve de l'antériorité pour se défendre en cas de litige.

Mais il existe aussi des désavantages :

- Révèle au grand jour des informations importantes sur l'invention, qui peuvent par la suite être utilisées par des concurrents.
- Elle peut limiter les options de commercialisation de l'invention.
- Elle peut rendre plus difficile l'obtention d'un brevet étendu dans d'autres pays (ce qui nous importe peu, car l'on vise le marché Suisse).

Il est donc important de consulter un conseil juridique avant de prendre une décision sur le dépôt d'une publication défensive.

3.4 Liberté d'exploitation - Titularité

Notre plateforme de recrutement inclut des technologies avancées telles que l'analyse de données, l'apprentissage automatique, la mise en correspondance des profils et l'analyse des compétences. Elle gère également les données des candidats, comme les CV et l'évaluation des compétences.

Après avoir effectué des recherches sur des bases de données de brevets comme *Espacenet* et *Swissreg*, nous avons découvert que certains domaines techniques de notre plateforme sont couverts par des brevets valides ; en voici deux : • WO2019068253A1 MACHINE LEARNING SYSTEM FOR JOB APPLICANT RESUME SORTING • FR2808354A1 PROCÉDE DE MISE EN CONCORDANCE AUTOMATIQUE D'UNE OFFRE D'EMPLOI AVEC UNE PLURALITE DE PROFILS Ces brevets couvrent l'apprentissage automatique pour le tri des CV de candidats ainsi que la mise en correspondance automatique entre une offre d'emploi et des profils de candidat. Afin d'éviter une infraction de brevet, nous pouvons négocier une licence avec les propriétaires pour utiliser les éléments protégés dans notre service. Par ailleurs, si nous souhaitons développer notre projet en collaboration avec un partenaire académique ou un collaborateur externe, nous devons rédiger un contrat de partenariat. Ce contrat doit définir les droits d'utilisation, les obligations et les contributions de chacun-e des partenaires pour la publication ainsi que pour la divulgation du projet. En revanche, même si nous sommes seuls à avoir développé le projet et que nous allons fonder une start-up pour développer et commercialiser notre projet, nous ne pouvons pas nous assurer être les propriétaires exclusifs de l'invention car nous ne disposons pas de brevet.

3.5 Contrats et litiges en matière de Propriété intellectuelle

Il y a trois sortes de contrats les plus fréquents auxquels on peut s'intéresser pour que notre service remplisse et utilise les contraintes légales au mieux : les contrats de confidentialité, les contrats de collaboration/développement et les contrats de licence. Observons les donc ici et voyons qu'en est-il pour notre service :

Le contrat de confidentialité est un accord par lequel une ou plusieurs parties s'engagent à ne pas divulguer des informations considérées comme confidentielles. Il est souvent utilisé pour protéger les données sensibles, les secrets commerciaux ou les informations de propriété intellectuelle.

Ainsi, nous y mettrons les conditions cités en section 3.6 ; c'est-à-dire le fait que nous limiterons la conservation et le traitement de ces données pour les besoins de la procédure de recrutement, qui prendra en général **moins de 6 mois**, ainsi que le fait que les données sensibles directes seront exclusivement utilisées pour l'identification du candidat, mais n'entreront pas en ligne de compte comme critère de sélection. Le type des données que cela concernera sera donc toutes les données présentes sur le CV du.e la candidat.e. Nous mettrons également dans ce contrat le fait que la récolte de CV se fait uniquement à des fins d'entraînement de modèle et de sélection de candidat.e.s, ainsi que la liste des corps de métiers de l'entreprise qui auront le droit d'y toucher, c'est-à-dire uniquement ceux qui vont travailler avec.

Aussi, nous spécifierons dans ce contrat le fait que les données seront détruites après qu'un délai d'une année maximum ce soit écoulée depuis la récolte des données.

Le contrat de collaboration/développement est un accord par lequel les parties s'engagent à travailler ensemble pour atteindre un objectif commun, généralement la création ou l'amélioration d'un produit ou d'un service. Il définit les responsabilités de chaque partie, les délais de développement, les coûts et les modalités de paiement.

Nous mettrons dans ce contrat le fait que la collaboration et le développement pour la construction du modèle de ML se fera seulement à l'interne, et que nul autre que les personnes travaillant sur ces données ou les consommant n'y aura accès, c'est-à-dire les data engineer / scientists et les personnes du secteur RH.

Les contrats de licence est un accord par lequel une partie (le titulaire des droits) autorise une autre partie (le licencié) à utiliser un bien protégé par des droits de propriété intellectuelle, comme un brevet, une marque déposée, une œuvre protégée par le droit d'auteur, ou un logiciel, en échange d'une rémunération ou de certains engagements. Il définit les droits d'utilisation, les restrictions d'utilisation, les coûts et les modalités de paiement, les clauses de confidentialité, les garanties et les procédures de résolution des conflits.

Nous mettrons dans ce contrat le fait que les entreprises utilisant notre service n'ont pas le droit de le monétiser pour d'autres entreprises et que toute transaction financière pour l'utilisation de ce service se fera uniquement des entreprises l'utilisant à notre entreprise.

C'est aussi ici que l'on spécifiera que les personnes utilisant notre service auront donc accès aux CV sélectionnés et sont tenus de garder une confidentialité stricte les concernant ; seulement les personnes du secteur RH ont le droit de les consulter. On spécifiera ici aussi les modalités de paiements et les coûts, que l'on déterminera au préalable à l'aide d'une étude marketing.

À noter qu'il est également important de se rappeler que les lois sur la protection des données et les lois sur l'éthique de l'IA doivent être respectées dans le cadre de ces contrats.

3.6 Protection des données

Comme toute plateforme de Hiring, nous allons faire face à la problématique de traitement des données utilisateurs. Étant donnée que nous nous limitons pour l'instant au marché Suisse comme mentionné en section 2, nous nous appuyerons sur la loi LPD (Loi Fédérale

sur la Protection des Données). Celle-ci régit le traitement de ces données afin de protéger au mieux la personnalité et les droits fondamentaux des personnes qui font l'objet d'un traitement de données en Suisse. À noter que nous recueillerons certainement les données de personnes non Suisses ou frontalières, de ce fait nous nous alignerons également à la RGPD (Règlement général sur la protection des données)

Nous nous baserons donc sur les prescriptions de ces deux en matière collecte, conservation, exploitation, modification, communication, archivage ou destruction des données personnelles.

Tout d'abord, une donnée personnelle peut être définie comme toute information se rapportant à une personne physique susceptible d'être identifiée ou identifiable, directement ou indirectement. Durant le processus de recrutement via notre plateforme, les postulant.e.s nous mettrons à disposition - via leur CV - différentes données personnelles afin de permettre l'étude de leur dossier. Pour ce qui est des données personnelles que nous récoltons, on les catégorise en deux groupes :

- **directes** : nom, prénom, photo, statut social ou marital ;
- **indirectes** : adresse postale, adresse e-mail, numéro de téléphone, date de naissance, études, formation professionnelle, nationalité, genre, langues parlées, formulaire de candidature ainsi que d'autres informations relatives à la candidature. Y compris la lettre de motivation, le CV avec - les expériences professionnelles antérieures - les qualifications professionnelles et autres compétences pertinentes - les références.

Parmi ces données, aucune n'est à caractère sensible. Toutefois, afin de respecter au mieux la LPD et la RGPD, nous devons :

- Vérifier la véracité des données avant un quelconque traitement ;
- Recueillir uniquement les données utiles au processus ;
- Tenir des postulant.e.s informé.e.s sur l'usage qui sera fait de leurs données ;
- Leur donner la possibilité de demander des informations portants sur l'utilisation de leurs données afin de garantir toute transparence ;
- Leur donner la possibilité de modifier ou de supprimer leur candidature ;
- Limiter les accès aux données personnelles récoltées : seules les personnes clés dans le processus pourront y accéder, entre autres les RH ;
- Les candidat.e.s seront immédiatement informé.e.s en cas de violation de leur données.

La LPD ne prévoit pas de délai de conservation spécifiques pour les données ; cette durée est définie au cas par cas. Afin d'éviter d'éventuelles violations suite à une durée de

conservation trop longue, nous limiterons la conservation et le traitement de ces données pour les besoins de la procédure de recrutement, qui prendra en général **moins de 6 mois**. Les données sensibles directes seront exclusivement utilisées pour l'identification du candidat, mais n'entreront pas en ligne de compte comme critère de sélection ou non - comme discuté d'ailleurs également en analyse des aspects éthiques de notre service, en section 4 de ce rapport.

En déposant sa candidature, le.la candidat.e à un emploi déclare avoir lu et compris notre politique en matière de confidentialité des données, laquelle expose la manière dont nous recueillons, traitons et utilisons ses données personnelles.

4 Aspects éthiques

Les aspects éthiques de ce projet soulèvent plusieurs questions quant au traitement des futur.e.s employé.e.s, à leurs sentiments et les biais inhérents au machine learning.

4.1 Biais du machine learning

Le machine learning sur des CV peut amener à un modèle qui apprend des données biaisées, qui peut provoquer du racisme, du sexisme, ou autres discriminations sociales. En effet, on constate par exemple aujourd'hui que plus d'hommes ont des postes à hautes responsabilités que de femmes, ou encore que plus d'hommes sont dans les secteurs de l'ingénierie, comme plus de femmes sont dans le secteur de la santé pour des métiers tels qu'infirmière, ou encore sage-femme.

Ainsi, il est important que notre modèle entraîné comporte le moins de biais possible pour être éthiquement le plus correct possible. La solution que nous proposons ici est de supprimer au possible les données qui peuvent être à caractère discriminant telles que le sexe du.de la candidat.e et son nom et prénom. On aimerait en effet ainsi éviter qu'un dataset d'entraînement déséquilibré au niveau de ces attributs engendre une discrimination sociale par le modèle.

Si les biais cités ci-dessus sont correctement évités, on pourra ainsi même espérer avoir moins de discriminations sociales faites par notre modèle de sélection de CV via l'IA que par certain.e.s recruteur.euse.s actuel.le.s.

4.2 Position du.de la futur.e employé.e

Le.la potentiel.le futur.e employé.e se fait ainsi, via ce service que nous proposons, juger par une machine. Ce fait peut entraîner un sentiment de perte d'humanité et peut également distancer la personne de l'entreprise à laquelle iel souhaite postuler, pour même risquer de la disculper de postuler. Le même problème d'éthique se pose d'ailleurs du côté des employé.e.s actuel.le.s, qui pourraient ainsi voir leur entreprise s'écarter de leurs valeurs.

Pour combattre ces sentiments néfastes, nous proposons d'être complètement transparents par rapport au modèle de machine learning que nous utiliserons ; typiquement en faisant signer au.à la CEO un document accessible par tou.te.s attestant le fait que les données pouvant provoquer des biais de sélection - tels que le sexe, nom, prénom comme cités précédemment - ne sont pas prises en compte pour la construction de notre modèle de classification.

5 Conclusion

En conclusion, les aspects éthiques et légaux sont importants à considérer lors de la création et de l'utilisation d'une plateforme de recrutement. Les différentes pratiques de recrutement éthiques étudiées permettent d'assurer l'égalité des chances, la transparence et la protection de la vie privée des candidat-e-s. Ceci permet de renforcer la confiance de nos client-e-s. Par ailleurs l'étude des aspects légaux nous a permis de garantir une conformité aux lois, notamment sur la confidentialité et la protection des données. Mais également sur l'utilisation des aspects techniques déjà breveté. Cela permet d'éviter les risques potentielles liés à l'infraction des brevets et protège notre propriété et notre fonds de commerce.