

Utiliser la communauté Open Law pour constituer une ressource commune partagée susceptible de faire référence

ARGUMENTAIRE ET ÉLÉMENTS DE CONTEXTE

Le potentiel que représente pour le droit "l'Intelligence artificielle" et les technologies associées (notamment *machine* et *deep learning*) est au cœur des préoccupations de la communauté des juristes. Pourtant, les applications sont aujourd'hui limitées, les professionnels de la legaltech manquant de données propres à la langue française et au langage juridique pour alimenter et faire progresser leurs algorithmes.

Cette absence de données d'apprentissage ne permet pas non plus d'aborder la question des dangers éthiques liés au développement des outils d'intelligence artificielle.

Ce constat est également celui de la CNIL . Cette dernière propose dans le cadre de son cycle « éthique et algorithmes » des organismes/associations/universités de travailler sur la notion d'éthique des algorithmes, de biais induits, de responsabilité... dans leur domaine. Elle rassemblera les restitutions de ces travaux à la rentrée, pour une publication en novembre. Avec son programme de constitution de datasets, Open Law fait partie des 30 partenaires qui ont répondu à cet appel.

C'est la raison pour laquelle l'association Open Law Le droit ouvert, avec le soutien de la CNIL dans le cadre de sa mission "Ethique et Algorithme", a décidé de solliciter sa communauté pour participer à un cycle exploratoire sur les données d'apprentissage, visant à :*

- **Démontrer la faisabilité et l'utilité** de constituer des jeux de données d'entraînement de qualité ;
- **Explorer des méthodologies de création** de données d'apprentissage ;
- **Documenter la démarche afin de lui conférer un caractère répliquable** pour d'autres jeux de données.

LE PROGRAMME EN BREF

Nous cherchons à travers un cas d'usage choisi collectivement, à démontrer la faisabilité et l'utilité de constituer des jeux de données d'entraînement de qualité. Le(s) jeu(x) de données produit dans le cadre de ce cycle exploratoire seront par la suite exploitables par tous (mise à disposition dans un format ouvert).

LE CAS D'USAGE RETENU

Le cas d'usage retenu pour ce cycle exploratoire est l'identification des différentes parties d'une décision de justice (parties, faits, procédure, exposé des motifs, dispositif...).

Ce zonage des décisions pourra permettre la création de nombreux outils :

- aide à la lecture des décisions,
- création de sommaires,
- identification des zones dans un but d'extraction d'informations précises...

La méthodologie issue de ce cas d'usage pourrait également servir à d'autres types de données, pour lesquelles le zonage est également nécessaire :

- textes,
- conventions collectives,
- contrats,
- appels d'offres...

CALENDRIER ET METHODOLOGIE DU PROGRAMME

ETAPE 1 - Avril - 2017

Définition du besoin - création du modèle d'annotation : en réunissant autour de la table les porteurs de technologies et les propriétaires de données, identification du cas d'usage retenu pour le projet et création du modèle d'annotation associé.

ETAPE 2 - Mai 2017

Mise en place et test de l'annotateur : à partir d'outil open source gratuit pour permettre la répliquabilité de la méthodologie

ETAPE 3 - Mai - Juillet 2017

Ateliers de crowdsourcing et formation machine learning : atelier publics et hébergés par des acteurs du monde du droit

ETAPE 4 - juillet-septembre 2017

Évaluation du datasets - corrections éventuelles - grâce à la constitution d'un comité éthique composé de juristes, statisticiens et développeurs

ETAPE 5 - sept - déc 2017

Utilisation des données, preuve de l'exploitabilité, restitution dans le cadre du Paris Open Source Summit (6 & 7 déc 2017)

POUR RÉUSSIR, NOUS AVONS BESOIN DE VOUS !

Conformément à l'ADN d'Open Law, ce programme se veut **le plus inclusif possible** : nous cherchons donc à réunir **tous les acteurs intéressés par la thématique du machine learning appliqué au Droit**.

Nous comptons sur ces derniers pour pleinement **contribuer à la définition de nos objectifs et des moyens nécessaires pour y parvenir** (évaluation du datasets, bonnes pratiques de machine learning, spécialistes des questions éthiques, porteurs de cas d'usage, etc).

QUI BÉNÉFICIERA DE NOS TRAVAUX ?

Les porteurs de technologies innovantes en droit qui disposeront d'un datasets de bonne qualité pour entraîner leur algorithme ;

Les producteurs de données qui bénéficieront d'une méthodologie simple pour créer les conditions d'un enrichissement collaboratif de leurs bases.

La communauté machine learning dans son ensemble, qui pourra s'appuyer sur datasets de référence en langue française, dont les applications pourraient s'étendre bien au-delà du monde du droit.

Le justiciable, qui a tout a gagné d'un climat général d'innovation favorable à une meilleure lisibilité et prévisibilité du droit

Les praticiens du droit, qui profiteront à terme en toute confiance de nouveaux outils métiers reposant sur l'intelligence artificielle

PRÉSENTATION RAPIDE

Open Law est une association et une communauté engagées dans la co-crédation de projets numériques au service de l'ouverture du droit. L'association intervient ainsi de façon prospective dans l'exploration de territoires inconnus et comme accélérateur d'innovation via l'organisation de programmes de co-crédation numériques courant sur des périodes de trois à six mois. Espace de confiance propice au dialogue, l'association cherche à éviter les blocages, anticiper les résistances, décloisonner les métiers, et montrer la faisabilité de l'ouverture et l'intérêt de la filière économique associée.

NOS VALEURS

Tous les projets conduits par l'association Open Law reposent sur les valeurs d'ouverture et de collaboration, tant dans la gouvernance mise en place que dans les méthodologie de réalisation et les modalités de diffusion de ces livrables. Tous les participants à nos programmes s'engagent à respecter la Charte Open Law, qui figure en annexe de ce document.

NOTRE DEMARCHE

Plus qu'une implication financière et technique, c'est l'investissement humain qui est fondamental pour intégrer cette communauté d'innovateurs et profiter pleinement des productions de nos projets. De cette manière, chacun bénéficie de l'effet de réseau et d'entraînement propre à la logique collective dans un environnement neutre propice au partage, à l'innovation, et à l'émergence de nouveaux usages. A notre niveau, nous souhaitons ainsi favoriser une dynamique positive susceptible d'apporter une première réponse collective aux enjeux contemporains, produit incrémentale de la créativité et de l'expertise de tous.

Le rôle d'Open Law ne consiste donc pas à accompagner directement des projets à vocation commerciale, mais à servir une communauté qui cherche à conduire collectivement les changements nécessaires dans le monde du droit.

POUR ATTEINDRE SES OBJECTIFS, L'association Open Law* Le droit ouvert :

Impulse et pilote des projets engageant des participants volontaires sur des réalisations concrètes (« livrables ») pour prototyper de nouveaux produits et services (bases de données, normes, inventaires, logiciels...), en mode « do-tank »

Met à la disposition de tous son ingénierie de collaboration : des méthodes et outils collaboratifs (hackathons, plateformes numériques, outils de consultation et de vote...), ainsi qu'une démarche dans laquelle il n'y a pas de « sachants » mais des « explorateurs »

Active sa communauté et recherche des partenaires en vue de monter des équipes pluridisciplinaires et compétentes, propose une assistance logistique dans la recherche de lieux de réunion, ainsi qu'une assistance de veille documentaire, et met à disposition des ressources techniques si besoin.

Applique une philosophie stricte d'ouverture et de partage de ses travaux, que ce soit au niveau des méthodes comme des réalisations qui sont mises à disposition de toutes les personnes intéressées (open data/open source) dans une bibliothèque de « communs »