

**Restitution CNIL - Programme Open Law
IA & droit - données d'apprentissage**

Le potentiel que représente l'intelligence artificielle pour le droit est au cœur des préoccupations de la communauté des juristes. Pourtant, les applications sont aujourd'hui limitées, car les professionnels de la *legaltech* manquent de données propres à la langue française et au langage juridique, données indispensables pour alimenter et faire progresser leurs algorithmes. Cette absence de données d'apprentissage de qualité pose aussi (et surtout) la question des dangers éthiques liés au développement des outils d'intelligence artificielle.

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, leurs enjeux éthiques et les méthodologies à appliquer pour y répondre.

L'association Open Law* Le Droit Ouvert

L'association Open Law*, Le Droit Ouvert est un espace de travail et d'expérimentation pour l'ensemble des acteurs du monde du droit prêts à innover dans un mode collaboratif et ouvert. Open Law se présente également comme un espace d'exploration de problématiques émergentes, telles que la Blockchain et les Smart Contrat, le Legal Design ou encore l'Open Gov.

Toutes ses actions aboutissent à la production de ressources communes ouvertes (référentiels, ontologies, livres blancs, modèles de statuts, cahiers des charges, bases de données, développements Open Source, etc.). Aujourd'hui, l'association Open Law réunit des représentants de l'ensemble des acteurs de l'information juridique et, plus largement, du monde du droit. Sa communauté rassemble environ 2 000 professionnels invités à interagir lors de plus de 100 temps forts depuis sa création.

1/ Présentation du projet

1.1/ Apprentissage automatique et données d'apprentissage

L'apprentissage automatique (ou *machine learning*) révolutionne la façon de concevoir les outils d'intelligence artificielle, en prenant le contre-pied des systèmes à base de règles (*rule-based*) utilisés communément dans le monde juridique. Un système à base de règle consiste à élaborer des modèles qu'on applique ensuite aux données. L'apprentissage automatique supervisé procède à l'inverse : l'homme fournit à l'algorithme des exemples de ce qu'il doit réaliser et l'algorithme produit lui-même le modèle en généralisant ces exemples. Par exemple, pour une tâche consistant à identifier dans des décisions de justice les prénoms et noms cités, on fournira des exemples de décisions dans lesquelles on aura « marqué » les prénoms et noms ; et l'algorithme utilisera ces exemples pour déterminer ce qui caractérise un nom propre (présence d'une majuscule en début de mot, présence de « M. » avant le mot, présence d'un verbe après le mot...).

Les avantages de cette technologie ne sont plus à démontrer. Cependant, cette approche comporte quelques risques : manque de transparence, amplification de biais... Pour ne s'intéresser ici qu'aux données d'apprentissage (les enjeux éthiques liés au choix des algorithmes et à leur auditabilité font l'objet de nombreuses études par ailleurs), leur importance est capitale. En effet, si les données fournies à l'algorithme ne sont pas cohérentes et de qualité (quantité insuffisante, mauvaise information, mauvaise représentation du réel), alors le résultat en sortie ne le sera pas non plus.

1.2/ Le projet IA & droit - données d'apprentissage

C'est dans ce contexte que 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 constituer un jeu de données utiles pour les outils de justice prédictive et ce faisant, informer sur les enjeux éthiques des données d'apprentissage et explorer des méthodologies de constitution de données d'apprentissage. L'objectif est, au travers d'un cas d'usage choisi collectivement, de démontrer la faisabilité et l'utilité de constituer des jeux de données d'entraînement de qualité.

Concrètement, le groupe de travail a sélectionné un cas d'usage utile pour l'ensemble de la communauté - le zonage des décisions de cours d'appel (voir détail ci-dessous). A partir de ce cas d'usage, il a été décidé de constituer un jeu de données d'apprentissage de plusieurs centaines de décisions. Un plan d'annotation a été établi en collaboration avec la Cour de cassation, ainsi que des instructions à destination des annotateurs. Un outil d'annotation *open source* a été mis en place. Des juristes de divers horizons (juristes d'entreprise, avocats, magistrats...) ont ensuite procédé au zonage des décisions au travers d'ateliers d'annotation.

Au cours de la constitution du jeu de données, un contrôle de la qualité des annotations a été appliqué. Le jeu de données produit dans le cadre de ce cycle exploratoire sera mis à disposition de tous dans un format ouvert. Une restitution complète de la méthodologie et des résultats de l'étude sera effectuée lors du Village de la legaltech le 7 décembre 2017. Le même jour, un événement public visant à promouvoir le jeu de données et en favoriser l'exploitation sera organisé.

1.3/ Le zonage des décisions de justice

Le cas d'usage retenu lors du programme est l'identification des différentes zones d'une décision de justice (présentation des parties, avocats ou magistrats, exposé des faits et de la procédure, exposé de la règle de droit applicable, argumentation du juge, solution...).

Le cas d'usage s'est imposé de lui-même. En effet, on voit – au sein de la communauté Open Law ou, plus largement, dans le monde de la *legaltech* – naître un grand nombre de projets se revendiquant de la justice prédictive et promettant, pour un contentieux donné, de prédire le montant alloué (calcul de risque), d'identifier les meilleurs arguments (aide à la construction de dossier), ou encore de rapprocher des décisions entre elles sur la base du fondement légal ou des arguments qui y sont développés (aide à la recherche). Toutes ces offres nouvelles ont pour point commun, outre d'éveiller un fort intérêt de leur public cible, de revendiquer l'utilisation de l'apprentissage automatique : extraction de montants, de concepts, de références légales... Mais ces projets, même de bonne qualité, ne sont aujourd'hui pas encore fiables pour deux raisons principales.

Ils ne sont pas fiables d'abord car nous ne disposons pas des données de première instance. Les extractions faites portent donc sur une petite portion du contentieux, et ne sont aucunement représentatives de ce qui est la réalité de la jurisprudence française. Avec l'adoption de la loi pour une République numérique du 7 octobre 2016, cette question est virtuellement résolue, puisque les décisions de première instance devront à l'avenir être disponibles en *open data*.

En second lieu, le manque de fiabilité de ces projets réside dans le fait que les informations qui doivent être extraites des décisions (montants alloués, concepts déterminants, fondement légal de la décision...) doivent être isolées d'autres informations similaires mais non utiles (montants demandés, arguments soulevés mais non retenus, textes invoqués sans succès...). Une analyse fine des projets en cours montre que personne pour le moment n'a réussi à résoudre cette problématique. Face à cette problématique, plusieurs membres d'Open Law ont identifié la nécessité de procéder à un zonage fin des décisions de justice qui permettrait d'améliorer les extractions de montants, concepts, fondements

légaux ou autres en les réalisant dans une zone précise. Des tentatives ont été faites pour identifier ces zones par règles, mais sans résultat satisfaisant, car la rédaction par les juges est très hétérogène. L'apprentissage automatique pourra résoudre cette difficulté.

2/ Les enseignements à mi-parcours

A l'heure où nous écrivons ces lignes, le projet est encore en cours. Une restitution complète présentera les résultats en décembre 2017. Voici néanmoins quelques premières conclusions.

2.1/ Bilan du zonage des décisions

Ses succès, tout d'abord : le jeu de données sera réalisé dans le temps imparti, et mis à disposition de tous. Le jeu n'est pas encore livré que des premières utilisations sont déjà prévues par les membres Open Law participant au projet. Il y a une véritable attente autour de ce projet.

Ses difficultés, ensuite : nous avons réalisé un plan d'annotation trop complexe par rapport à notre objectif de départ, qui supposait de recourir au *crowdsourcing*. Nous avons donc, d'une part, allégé le plan d'annotation et, d'autre part, décidé de limiter le nombre d'annotateurs et de les accompagner. De même, nous nous sommes limités au zonage des décisions d'appel en matière judiciaire, afin de respecter les délais que nous nous étions fixés. Enfin, nous avons revu un peu à la baisse nos ambitions en terme de volume de décisions annotées.

Ses surprises, enfin : une première analyse des données annotées nous indique que les catégories qui seront les plus faciles à entraîner (la phase d'apprentissage n'est pas encore commencée à l'heure où nous écrivons ces lignes) ne sont pas celles qu'on attendait. Ainsi, nous pensions que l'identification des parties, avocats, magistrats de la décision serait aisée. En réalité, ces informations sont bien souvent mélangées et les zones difficiles à dissocier. De même, nous pensions qu'il serait possible de rapprocher grâce au zonage la solution (dans le dispositif) et l'argumentation du juge correspondante (motif). Cela ne sera *a priori* pas du tout possible. A l'inverse, les références de la décision attaquée sont formulées de façons très variées mais toujours isolées du reste de la décision, l'apprentissage devrait donc être relativement facile sur cette zone. De même, l'identification de la règle de droit devrait être faisable,

même lorsque la règle de droit énoncée ne provient pas d'un texte clairement cité ou d'une décision de justice.

De façon plus étonnante encore, ce projet de zonage permettra de développer des outils bien au-delà des seuls outils dits de justice prédictive. Ainsi, des projets de recherche dans les décisions ou encore un projet d'analyse qualitative du travail de rédaction des magistrats sont envisagés. C'est d'ailleurs pour favoriser ces exploitations « nouvelles » du jeu de données qu'un hackathon est organisé le 7 décembre prochain lors du Village de la legaltech.

Enfin, ce projet aura grandement contribué à la formation des acteurs de la communauté juridique sur cette technologie qu'est le *machine learning* et sur ses usages possibles, et à démystifier la fameuse « justice prédictive ». Lors des ateliers d'annotation, une formation au *machine learning* a été créée et plébiscitée. Au point de faire des ateliers d'annotation sans la moindre annotation...

2.2/ Evaluation et correction de la qualité du jeu de données

Au cours du projet, nous avons rencontré plusieurs problématiques liées à la qualité de l'annotation : l'algorithme cherchant à définir le modèle sous-tendant les exemples qui lui sont soumis, il est absolument indispensable de fournir des données soit en très grand nombre, soit d'excellente qualité, cohérentes entre elles et représentatives de l'ensemble des données. En matière juridique, nous ne disposons pas d'un nombre suffisant de données d'apprentissage pour que les erreurs qui s'y glissent soient quantité négligeable, il est donc indispensable de travailler sur la qualité du jeu de données.

En premier lieu, il convient de travailler la qualité des annotations elles-mêmes. Pour ce faire, il faut s'assurer de l'adéquation entre la tâche confiée aux annotateurs, leurs compétences et disponibilités et l'outil qui est mis à leur disposition. Par exemple, un projet prévoyant une tâche longue et pénible et nécessitant un haut niveau d'expertise sur un outil difficile à manipuler sera voué

à l'échec, car il sera impossible de trouver les utilisateurs experts, disponibles et motivés pour la tâche.

En fonction du projet envisagé et de ses contraintes, il faudra donc trouver le meilleur équilibre entre :

- **le niveau de complexité du plan d'annotation** : le plus simple possible, pour que la tâche puisse être faite sans trop de réflexion, mais néanmoins utile ;
- **le niveau d'expertise et la disponibilité des annotateurs** : en fonction du projet, il faudra recourir à des annotateurs plus ou moins spécialisés - ainsi, par exemple, on constate pour notre projet que les annotateurs les plus efficaces sont ceux qui ont l'habitude de passer en revue des arrêts sans forcément les lire dans le détail (veilleurs, greffiers, services sources des éditeurs juridiques, magistrats...) et qu'à l'inverse la tâche est très difficile pour des avocats, habitués à analyser le fond des arrêts et qui n'arrivent pas à les survoler ;
- **la qualité de l'outil d'annotation** : outils adaptés à la tâche (classification, reconnaissance d'entité...), outils permettant d'intégrer des règles pour effectuer des pré-annotations que les annotateurs devront simplement vérifier, outils combinant annotation et entraînement du modèle, pour alléger le travail de l'annotateur au fur et à mesure de l'avancement du projet... ;
- **la possibilité d'automatiser une partie des annotations** par des règles lorsque cela est possible, en faisant attention bien sûr à en contrôler le résultat.

A noter : on pourra subdiviser le projet en plusieurs sous-tâches de complexités différentes.

Les méthodes d'évaluation de la qualité des annotations sont relativement variées et plutôt bien documentées. Une évaluation manuelle est évidemment toujours possible, mais on peut aussi automatiser les sondages via des techniques de comparaison des annotations, entre les annotateurs (calcul de *l'interagreement*) ou par rapport à un référentiel pré-établi et validé comme correct (que nous appellerons « *gold dataset* »). Nous avons tenté chacune de ces méthodes, et les analyses sont en cours.

Enfin, un second critère de qualité est celui de la qualité du jeu de données dans son ensemble. Si le jeu de données d'apprentissage ne représente pas correctement l'ensemble des cas de figure présents dans les données brutes à enrichir, alors l'algorithme risque de d'ignorer certaines informations. Il convient donc de s'assurer de bien connaître le jeu de données global (statistiques de représentation géographique, temporelle, matérielle...) parallèlement au travail de constitution du jeu de données d'apprentissage, afin de pouvoir vérifier la représentativité de ce dernier.